



**Universidad Autónoma de Querétaro**

**Facultad de Informática**

Sistema predictivo de rendimiento escolar de alumnos universitarios  
de primer ingreso

**Tesis**

Que como parte de los requisitos  
para obtener el Grado de

**Doctor en Innovación en Tecnología Educativa**

Presenta

**Enrique Ayala Franco**

Dirigido por:

Dra. Rocío Edith López Martínez

Co-Director:

Dr. Víctor Hugo Menéndez Domínguez

Querétaro, Qro. a 5 de julio de 2022



**Universidad Autónoma de Querétaro**  
**Facultad de Informática**  
**Doctorado en Innovación en Tecnología Educativa**

Sistema predictivo de rendimiento escolar de alumnos universitarios de primer ingreso

**Tesis**

Que como parte de los requisitos para obtener el Grado  
Doctor en Innovación en Tecnología Educativa

Presenta:  
Enrique Ayala Franco

Dirigido por:  
Dra. Rocío Edith López Martínez

Co-dirigido por:  
Dr. Víctor Hugo Menéndez Domínguez

Dra. Rocío Edith López Martínez  
Presidente

Dr. Víctor Hugo Menéndez Domínguez  
Secretario

Dra. Ma. Teresa García Ramírez  
Vocal

Dr. Ricardo Chaparro Sánchez  
Suplente

Dra. Diana Margarita Córdova Esparza  
Suplente

Centro Universitario, Querétaro, Qro.  
Julio de 2022  
México

## DEDICATORIAS

A Dios, quien nos ha permitido cumplir nuestros sueños y anhelos, bajo su manto protector.

A mi esposa Yazz, quien me ha brindado su apoyo incondicional todos estos años y siempre ha sabido motivarnos para alcanzar nuestras metas.

A mis hijos Juan Enrique y Jesús Virgilio, quienes con su alegría nos impulsan siempre a seguir adelante.

A mi familia, en especial a mi madre Ángela y mi suegra Conchita, pues estos momentos sólo tienen sentido si los compartimos con quienes siempre nos han querido y apoyado.

## **AGRADECIMIENTOS**

A mi directora de tesis, Dra. Rocío Edith López Martínez y mi codirector, Dr. Víctor Hugo Menéndez Domínguez, quienes con su experiencia y conocimientos supieron orientar con precisión el desarrollo de las actividades del proyecto de tesis, además de apoyarme y motivarme en todo momento.

Al Programa para el Desarrollo Profesional Docente (PRODEP), por la beca de Apoyo para estudios de posgrado de alta calidad, de la Subsecretaría de Educación Superior de la Secretaría de Educación Pública.

A la Facultad de Informática de la Universidad Autónoma de Querétaro, y a la coordinadora del DITE, Dra. María Teresa García Ramírez, quién nos recibió y acompañó durante todos los estudios, siempre con una sonrisa en su rostro.

A los integrantes del sínodo por sus revisiones y comentarios siempre valiosos.

# ÍNDICE

<i>DEDICATORIAS</i> .....	<i>i</i>
<i>AGRADECIMIENTOS</i> .....	<i>ii</i>
<i>ÍNDICE</i> .....	<i>iii</i>
<i>ÍNDICE DE TABLAS</i> .....	<i>vii</i>
<i>ÍNDICE DE FIGURAS</i> .....	<i>viii</i>
<i>RESUMEN</i> .....	<i>x</i>
<i>ABSTRACT</i> .....	<i>xi</i>
1. <i>INTRODUCCIÓN</i> .....	1
2. <i>ANTECEDENTES</i> .....	7
2.1 Categorías del Análisis Documental.....	7
2.2 Determinantes del Rendimiento Escolar.....	9
2.3 Evaluación del Rendimiento Académico.....	11
2.4 Tecnologías en el Contexto Educativo .....	12
2.5 Minería de Datos Educativa y Rendimiento Académico .....	15
2.6 Estrategias de Intervención Educativa.....	21
2.7 Líneas de Investigación Emergentes.....	25
3. <i>MARCO CONCEPTUAL Y TEÓRICO</i> .....	30
3.1 Perspectiva del Desarrollo del Estudio .....	30
3.2 Rendimiento Académico.....	31
3.2.1 Influencia de Variables Socioemocionales.....	32
3.2.2 Medición del Rendimiento Académico .....	33
3.2.3 Bajo Rendimiento y Riesgo Académico .....	35
3.3 Transición Universitaria .....	36
3.4 Intervención Educativa .....	37
3.5 Educación Mediada por Tecnología .....	38
3.5.1 Ambientes Virtuales de Aprendizaje .....	39
3.6 Diseño Instruccional .....	41
3.6.1 Perspectiva Teórica del Diseño Instruccional .....	41
3.6.2 Modelo de Diseño Instruccional Mediado con Tecnología .....	43

3.7	Minería de Datos Educativos.....	44
3.7.1	Tareas de la Minería de Datos Educativos .....	45
3.8	Instrumentos Teóricos de Análisis de Datos.....	47
3.8.1	Técnicas de Minería de Datos Educativos .....	47
3.8.2	Validación de Modelos Predictivos.....	50
4.	<i>METODOLOGÍA</i> .....	53
4.1	Determinación de Hipótesis y Premisa .....	53
4.2	Objetivos del estudio .....	54
4.2.1	Objetivo General .....	54
4.2.2	Objetivos Específicos.....	54
4.3	Tipo de Investigación.....	54
4.3.1	El Enfoque de la Metodología Mixta.....	55
4.4	Proceso de Investigación-Acción.....	57
4.5	Proceso de la Minería de Datos Educativos .....	58
4.6	Alcance de la Investigación .....	61
4.7	Población y Muestra .....	61
4.8	Estrategia de Recolección y Análisis de Datos.....	62
4.9	Recursos Materiales y Humanos .....	63
4.10	Procedimientos .....	64
5.	<i>PROCESO DE MINERÍA DE DATOS EDUCATIVOS</i> .....	68
5.1	Extracción de Datos.....	68
5.2	Procesamiento previo .....	70
5.3	Aplicación de Técnicas de MDE .....	73
5.3.1	Herramientas de la MDE.....	73
5.3.2	Selección de Variables Significativas.....	74
5.3.3	Aplicación de Técnicas de Clasificación .....	76
5.4	Generación de Modelos Predictivos .....	77
5.4.1	Algoritmo J48 .....	77
5.4.2	Algoritmo RandomForest .....	80
5.4.3	Algoritmo LMT .....	80
5.4.4	Algoritmo Logistic.....	82
5.4.5	Algoritmo MultilayerPerceptron .....	84
5.5	Validación del Modelo.....	86
5.6	Interpretación de los Modelos Resultantes.....	86
5.7	Conclusiones de la Etapa de MDE .....	90

6.	<i>PROCESO DE IMPLEMENTACIÓN DEL SISTEMA PREDICTIVO</i> .....	93
6.1	Componentes y Tecnologías Utilizadas.....	93
6.2	Arquitectura del Sistema.....	94
6.3	Recuperación de Datos de Encuestas desde Moodle .....	95
6.4	Selección de Modelos Predictivos .....	97
6.5	Predicción de Riesgo.....	97
6.6	Comunicación de los Resultados .....	99
6.7	Módulo de Seguimiento a Cursos.....	100
6.8	Comentarios Finales.....	102
7.	<i>PROCESO DE INVESTIGACIÓN-ACCIÓN</i> .....	104
7.1	Análisis Diagnóstico.....	104
7.1.1	Enfoque Metodológico para el Diagnóstico.....	105
7.1.2	Resultados del Diagnóstico.....	106
7.1.3	Discusión de los Resultados del Diagnóstico.....	114
7.1.4	Conclusiones del Análisis Diagnóstico.....	116
7.2	Planeación de la Intervención Educativa .....	117
7.2.1	Prioridades.....	118
7.2.2	Beneficios Esperados de la Intervención .....	118
7.2.3	Metas de la Intervención .....	119
7.2.4	Situación Inicial .....	120
7.2.5	Líneas de Acción.....	122
7.3	Implementación del Plan .....	123
7.3.1	Detección de Alumnos en Riesgo .....	123
7.3.2	Seguimiento de Cursos.....	124
7.3.3	Diseño e Impartición de Curso para el Desarrollo de Habilidades Académicas .....	126
7.4	Análisis Cuantitativo de la Intervención .....	129
7.4.1	Medición del Efecto de Intervención en el Perfil Autodirigido.....	129
7.4.2	Resultados de la Encuesta CIPA+ .....	131
7.5	Análisis Cualitativo de la Intervención .....	135
7.5.1	Observaciones del Personal Académico Involucrado .....	136
7.5.2	Percepción de los Alumnos.....	138
8.	<i>INTEGRACIÓN DE RESULTADOS Y DISCUSIÓN</i> .....	147
8.1	Resultados de la Etapa de Minería de Datos .....	147
8.2	Resultados Generados por el Sistema Predictivo .....	149
8.2.1	Predicción .....	149
8.2.2	Seguimiento .....	151

8.3	Resultados del Diagnóstico y la Planeación de la Intervención .....	154
8.4	Resultados de Aspectos Cuantitativos de la Intervención .....	155
8.4.1	Efectos en el Rendimiento Académico.....	155
8.4.2	Efectos en el Perfil Autodirigido .....	156
8.5	Resultados de Aspectos Cualitativos de la Intervención .....	158
8.6	Discusión .....	162
9.	<i>CONCLUSIONES</i> .....	166
10.	<i>REFERENCIAS</i> .....	172
11.	<i>ANEXOS</i> .....	180
	Anexo 1. Instrumento Diagnóstico.....	180
	Anexo 2. Planeación Didáctica del Curso .....	183
	Anexo 3. Características de los Recursos TIC Empleados .....	190
	Anexo 4. Códigos y Transcripciones del Análisis Cualitativo .....	192

## ÍNDICE DE TABLAS

Tabla 1 Sistema de categorías de análisis. ....	8
Tabla 2 Descripción del conjunto de variables inicial. ....	71
Tabla 3 Resultados de 3 métodos de selección de atributos. ....	74
Tabla 4 Resultados del cuarto método de selección de atributos. ....	75
Tabla 5 Reglas de conocimiento derivadas del árbol de decisión J48 para la clase RIESGO_P. ....	78
Tabla 6 Reglas de conocimiento derivadas del árbol de decisión J48 para la clase RIESGO_A. ....	79
Tabla 7 Resultados de la evaluación de los modelos predictivos. ....	86
Tabla 8 Principales apoyos requeridos por los estudiantes. ....	113
Tabla 9 Aspectos a mejorar en el entorno de la escuela. ....	114
Tabla 10 Líneas de acción para la intervención educativa. ....	122
Tabla 11 Distribución de frecuencias por nivel de autodirección integrado, antes de la intervención. ....	131
Tabla 12 Distribución de frecuencias por nivel de autodirección integrado, después de la intervención. ....	132
Tabla 13 Resultados de niveles de autodirección por componente, Grupo 1. ....	133
Tabla 14 Resultados de niveles de autodirección por componente, Grupo 2. ....	134
Tabla 15 Matriz de confusión de los resultados de la predicción. ....	150
Tabla 16 Tasas de acierto de la predicción y resultados reales. ....	150

## ÍNDICE DE FIGURAS

Figura 1 Caso de diseño de múltiples fases de la metodología mixta. ....	56
Figura 2 Fases del proceso de Investigación-Acción. ....	58
Figura 3 Proceso de la minería de datos educativa. ....	59
Figura 4 Diseño general del proyecto de investigación. ....	65
Figura 5 Tareas realizadas en la etapa de minería de datos educativos. ....	68
Figura 6 Árbol de modelo logístico (LMT) para la clase RIESGO_A. ....	81
Figura 7 Árbol de modelo logístico (LMT) para de la clase RIESGO_P. ....	82
Figura 8 Coeficientes del modelo logístico para la clase RIESGO_A. ....	83
Figura 9 Coeficientes del modelo logístico para la clase RIESGO_P. ....	84
Figura 10 Modelo parcial MultilayerPerceptron para la clase RIESGO_P. ....	85
Figura 11 Arquitectura del sistema predictivo. ....	95
Figura 12 Conexión inicial a la plataforma Moodle desde el sistema predictivo. ...	96
Figura 13 Selección de curso y encuesta Moodle para la recuperación de datos. ...	96
Figura 14 Selección de modelos predictivos a utilizar. ....	97
Figura 15 Resultado de la predicción mediante el algoritmo Logistic. ....	98
Figura 16 Mensajes con resultados de predicción enviados a los alumnos. ....	99
Figura 17 Restricción de acceso a recurso Moodle con base en calificación de una actividad. ....	100
Figura 18 Ítems de la encuesta de satisfacción para seguimiento periódico. ....	100
Figura 19 Módulo de seguimiento de satisfacción de los cursos. ....	101
Figura 20 Resultados comparativos globales de los cursos. ....	102
Figura 21 Pasos generales de la investigación-acción. ....	104
Figura 22 Resultados generales de promedios de calificaciones y materias reprobadas de las carreras de computación (2016-2018). ....	107
Figura 23 Resultados de promedios de calificaciones y materias reprobadas LIC (2016-2018). ....	107

Figura 24 Resultados de promedios de calificaciones y materias reprobadas LIS (2016-2018). .....	108
Figura 25 Resultados de promedios de calificaciones y materias reprobadas LCC (2016-2018). .....	108
Figura 26 Opinión de los estudiantes sobre antecedentes y expectativas personales. ....	110
Figura 27 Opinión de los estudiantes sobre situaciones personales y apoyos requeridos. ....	111
Figura 28 Opinión de los estudiantes sobre uso de tecnología educativa novedosa. ....	112
Figura 29 Opinión sobre el programa de tutorías institucional. ....	113
Figura 30 Elementos considerados en el diseño del curso. ....	127
Figura 31 Nube de códigos del primer análisis. ....	139
Figura 32 Códigos de la categoría actividades significativas .....	139
Figura 33 Modelo para el desarrollo y aplicación del sistema predictivo. ....	147
Figura 34 Resultados globales de satisfacción de los cursos. ....	152
Figura 35 Gráficas de resultados de satisfacción para Álgebra superior. ....	153
Figura 36 Comparación de promedios de calificaciones, históricas y grupo intervenido. ....	155
Figura 37 Comparación de materias reprobadas, históricas y grupo intervenido. ....	156
Figura 38 Distribución de frecuencias por niveles de autodirección, antes y después del curso. ....	157
Figura 39 Red semántica para la categoría aprendizajes favorecidos. ....	160

## RESUMEN

Los problemas derivados del bajo rendimiento académico, como el rezago y la deserción, son frecuentes en los primeros períodos de estudio universitario. En el caso de las carreras de computación de la Universidad Autónoma de Yucatán, el primer semestre es crucial, ya que hasta un 44% de los estudiantes llegan a reprobado 2 o más asignaturas. Ante la variedad de factores personales, sociales, académicos e institucionales que inciden en el bajo rendimiento escolar, es necesario obtener y analizar información para identificar posibles problemáticas de manera oportuna. El objetivo del proyecto fue desarrollar un sistema predictivo de rendimiento escolar de alumnos de primer ingreso a carreras del área de computación, utilizando modelos generados mediante minería de datos educativa; con el fin de facilitar la detección y atención de problemas asociados al bajo rendimiento y desplegar estrategias de intervención educativa. La metodología empleada sigue un enfoque de investigación mixto. En la primera etapa, con un enfoque cuantitativo y de diseño no-experimental, se utilizaron procedimientos de la minería de datos educativos para generar modelos predictivos eficientes de rendimiento escolar, que posteriormente fueron implementados en un sistema informático para la predicción de riesgo académico de nuevos estudiantes. En la segunda etapa, con orientación cualitativa, se aplicaron estrategias de intervención educativa dirigidas a la prevención y atención del bajo rendimiento. Destaca el diseño de un curso en modalidad e-learning para fomentar el desarrollo de habilidades de autodirección y la adaptación a la escuela. Se obtuvieron modelos predictivos de riesgo escolar y la identificación de los mejores atributos. La implementación y uso del sistema, generó información para aplicar estrategias de intervención más focalizadas, contribuyendo con el perfil de autodirección de los alumnos. Las opiniones de los involucrados evidencian mejoras en el rendimiento académico y en el desarrollo de sus habilidades académicas.

**(Palabras clave:** minería de datos educativos, riesgo académico, sistema predictivo, intervención educativa, autodirección.)

## ABSTRACT

Problems derived from poor academic performance, such as falling behind and dropping out, are frequent in the first periods of university study. In the case of computer careers at the Autonomous University of Yucatan, the first semester is crucial, since up to 44% of students fail 2 or more subjects. Given the variety of personal, social, academic and institutional factors that affect poor school performance, it is necessary to obtain and analyze information to identify possible problems in a timely manner. The objective of the project was to develop a predictive system for school performance of first-year students in computer science, using models generated by educational data mining; in order to facilitate the detection and attention of problems associated with poor performance and deploy educational intervention strategies. The methodology used follows a mixed research approach. In the first stage, with a quantitative approach and non-experimental design, educational data mining procedures were used to generate efficient predictive models of school performance, which were later implemented in a computer system for the prediction of academic risk of new students. In the second stage, with a qualitative orientation, educational intervention strategies aimed at preventing and addressing low performance were applied. The design of an e-learning course to promote the development of self-management skills and adaptation to school stands out. Predictive models of school risk and the identification of the best attributes were obtained. The implementation and use of the system generated information to apply more focused intervention strategies, contributing to the self-direction profile of the students. The opinions of those involved show improvements in academic performance and in the development of their academic skills.

**(Keywords:** educational data mining, academic risk, predictive system, educational intervention, self-direction.)

# 1. INTRODUCCIÓN

En este capítulo se presenta el contexto de la problemática relacionada con el bajo rendimiento académico y la necesidad de efectuar análisis de datos generados en entornos educativos, empleando nuevos enfoques metodológicos y herramientas de software basados en la minería de datos educativos y las analíticas del aprendizaje. Posteriormente, se hace el planteamiento formal del problema de investigación y la justificación para realizar el estudio. Finalmente, se describe la estructura y los apartados que conforman el documento de tesis.

A pesar de los múltiples estudios sobre rendimiento académico aún se siguen manifestando de forma preocupante las problemáticas relacionadas al mismo. En particular en México, de acuerdo con la Organización para la Cooperación y el Desarrollo Económicos (OECD por sus siglas en inglés), algunos de los principales problemas que presentan la mayoría de las instituciones educativas son: la deserción y el bajo rendimiento escolar; en los cuales confluyen factores de orden personal, familiar, de contextos educativos y laborales. Entre otros motivos, el abandono de la escuela se suma a las causas por las que, en 2017, en México, sólo el 26% de los adultos jóvenes tenían estudios de nivel secundario superior y el 23% un título de nivel terciario, aún muy por debajo de los promedios de la OCDE, del 41% y el 44%, respectivamente (OECD, 2018).

Ante este panorama, las instituciones educativas se han planteado elevar el rendimiento académico de sus estudiantes atendiendo algunos de sus factores determinantes; para lograr este propósito se ha tornado imprescindible contar con información que facilite el entendimiento de las causas que influyen en los resultados educativos y que revelen posibles problemáticas para intervenir con oportunidad y eficacia. Las fuentes de la información pueden ser muy diversas, sin embargo, el uso cada vez más extendido de sistemas informáticos ha generado una gran cantidad de bases de datos que se utilizan generalmente con fines

administrativos, pero que pudieran aprovecharse para explorar aspectos relacionados con la educación.

En la Universidad Autónoma de Yucatán (UADY), desde hace varios años, se utilizan sistemas de información para la gestión de datos escolares, por ejemplo, el denominado Sistema de Información y Control Escolar Institucional (SICEI). También se cuenta con plataformas para la impartición de cursos en modalidad virtual o en línea, además de otros sistemas para la evaluación docente y la evaluación curricular, cada uno de los cuales genera de manera rutinaria gran cantidad de datos en formato digital. Desafortunadamente este caudal de información no se ha empleado cabalmente para realizar investigaciones educativas u obtener información novedosa y relevante sobre los alumnos y sus procesos formativos de manera que se pudieran detectar situaciones problemáticas e intervenir con mayor oportunidad.

Por otro lado, en la Facultad de Matemáticas de la UADY operan seis programas de estudio, tres relacionados con las matemáticas y los otros tres concernientes a la computación. En general, se han tenido altos índices de reprobación en los programas del área de computación, ya que en promedio llegan a reprobado 44% o más de las materias en el primer año de estudios. Este arranque complicado de los estudiantes les genera varios obstáculos para el resto de la trayectoria escolar, por ejemplo, dificultades para inscribir materias por traslapes en horarios y excesiva carga de trabajo por presentar exámenes o llevar cursos de regularización; en consecuencia, muchas de las veces abandonan sus estudios o son dados de baja por exceder las oportunidades para acreditar sus materias.

Para apoyar a los estudiantes existe un programa institucional de tutorías en el cual profesores-tutores actúan como guías en el medio académico para orientarlos durante su trayectoria escolar; sin embargo, en el primer semestre no reciben de manera oportuna y sistematizada información sobre los perfiles de los estudiantes,

por lo cual es complicado hacer un diagnóstico preciso sobre las problemáticas y factores que pudieran incidir en el rendimiento académico.

A pesar de contar con sistemas de información institucionales, la carencia de mecanismos y procedimientos para extraer información relevante de la población estudiantil, principalmente en el primer semestre de estudios, afecta la toma de decisiones de los directivos y docentes que, al no contar con una visión más precisa de los alumnos que reciben y sus problemáticas, dejan de implementar acciones para la mejora educativa. Por los motivos expuestos, se puede ver la importancia de desarrollar y aplicar nuevas metodologías e instrumentos para el análisis de los datos disponibles en la Institución, para lograr un mejor entendimiento de los estudiantes, sus procesos formativos y las causas de su rendimiento académico, además de estrategias de atención y prevención.

Para el planteamiento formal de la problemática en estudio se formularon las siguientes preguntas de investigación:

- ¿Cómo entender, predecir el rendimiento escolar de los estudiantes de nuevo ingreso de la Facultad de Matemáticas de la UADY, utilizando herramientas de software?
- ¿Qué factores extraescolares o académicos determinan significativamente el rendimiento académico de los estudiantes de la Facultad de Matemáticas de la UADY?
- ¿Qué estrategias de intervención educativa son apropiadas para modificar, mejorar el rendimiento escolar de alumnos con antecedentes problemáticos en la Facultad de Matemáticas de la UADY?

Se pretende que la información, los modelos y sistemas generados durante el desarrollo de este trabajo, sirvan como instrumentos estratégicos para la toma de decisiones oportuna de tutores, docentes, directivos y alumnos; de manera que

impacte de manera positiva en el rendimiento académico de los estudiantes y se mitiguen otro tipo de problemáticas como el rezago y la deserción.

Dada la complejidad y multidimensionalidad del rendimiento académico, es improbable llevar a cabo estudios que mapeen cada una de las causas que influyen en el mismo, sin embargo, desde nuestro punto de vista, es factible encontrar los principales factores, tomando como insumo la información disponible del historial académico y datos extraescolares, para generar modelos con un aceptable nivel de precisión que se puedan utilizar para atisbar el futuro académico de los estudiantes y con ello, en la medida de lo posible, atender algunas de las más importantes causas del bajo rendimiento escolar en su contexto educativo.

Durante el desarrollo del proyecto de investigación, se considera el diseño y aplicación de estrategias de intervención educativa para atender las principales problemáticas detectadas en los estudiantes en riesgo, con énfasis en las carreras con mayores problemas de rendimiento académico. Las estrategias son aplicables a nivel de aula o de la escuela, y requieren la coordinación de actividades con directivos, responsables de programas y profesores-tutores para su diseño y aplicación. Con el desarrollo del proyecto, se espera que se reconozcan las bondades del empleo de nuevas metodologías y herramientas tecnológicas en el ámbito académico para analizar datos y obtener información de forma más eficiente y oportuna; y con ello se dé pie a su adopción en otros procesos educativos, promoviendo con ello la innovación y la mejora en la Institución.

Por otro lado, se ha observado que los factores detectados en ciertos programas de estudio pueden variar en relación con los encontrados en otros, es decir, las particularidades geográficas, socioeconómicas, temporales y específicas de la población estudiantil y las escuelas, pueden impedir la generalización de los resultados, por ello la necesidad de realizar estudios para analizar contextos determinados y concretos de las instituciones en la búsqueda de predictores de rendimiento académico (Vásquez, 2016). Es decir, se requieren estudios que

abonen al conocimiento de los determinantes del desempeño escolar de grupos de estudiantes específicos, en entornos académicos locales.

El documento de tesis está estructurado en apartados que dan cuenta de los elementos que delimitan y sustentan el proyecto de investigación realizado, acorde a la metodología científica.

Antecedentes y estado del arte, en donde se presentan los trabajos que han precedido a esta obra y que aportan elementos de referencia, contraste y análisis respecto a los temas de rendimiento académico, sistemas basados en minería de datos educativas para la predicción del desempeño escolar e implementación de intervenciones educativas.

Marco conceptual y teórico, que compendia la revisión de los principales referentes relacionados con el tema de estudio. Se establece la perspectiva desde la que se aborda el aprendizaje y a continuación se presentan las aproximaciones conceptuales sobre rendimiento académico, la transición universitaria, los procesos de intervención educativa, así como los principios sobre educación mediada por tecnología y particularmente, el uso de la minería de datos educativos como principal herramienta de análisis de los datos del estudio.

Metodología, donde se presentan la hipótesis, los supuestos de la investigación y los objetivos a lograr con el desarrollo de este trabajo; también se detalla el enfoque de trabajo empleado para el desarrollo del proyecto, en donde se sustenta la selección de la metodología mixta, la cual nos permite trabajar con varios enfoques de recolección y análisis de datos, y aporta una visión más amplia del fenómeno estudiado. En este apartado también se describe el proceso de la investigación-acción para el diseño de la intervención educativa y las etapas de la minería de datos educativos.

Acorde a la metodología, se desarrollan y presentan secciones con las etapas realizadas, en donde se describen las acciones, sus resultados y se presentan

conclusiones correspondientes a cada uno de los momentos. Se describe también el proceso de desarrollo e implementación del sistema. Los apartados finales de la tesis presentan una integración de resultados, en donde se comentan y discuten los hallazgos de las etapas previas, lo cual permite dimensionar el alcance del trabajo realizado al contrastar con los referentes teóricos y los trabajos previos relacionados con el objeto de estudio y los objetivos.

En la sección de Conclusiones se enlistan las aportaciones con las que este trabajo de investigación contribuye al conocimiento científico, haciendo las acotaciones pertinentes al señalar las limitaciones del estudio, y proponiendo nuevas líneas de investigación para profundizar o complementar el trabajo realizado.

## **2. ANTECEDENTES**

En este capítulo, tomando en cuenta los objetivos y preguntas de investigación, se presenta una revisión de estudios relacionados con: determinantes del rendimiento escolar, evaluación del rendimiento académico, minería de datos educativa y predicción del rendimiento académico, para identificar factores intervinientes y generar modelos predictivos; además de analizar trabajos relacionados con estrategias de intervención educativa, prioritariamente para estudiantes de primer año de instituciones educativas de nivel superior, que presenten alternativas de solución a las problemáticas derivadas del bajo rendimiento académico. Finalmente, se exploran trabajos sobre educación mediada por tecnología y el uso de ambientes virtuales de aprendizaje, como formas novedosas de apoyo a los procesos intervención y seguimiento del proceso de aprendizaje.

### **2.1 Categorías del Análisis Documental**

El proceso para realizar el análisis sigue un enfoque cualitativo e interpretativo, de tipo documental (Gómez et al., 2015). Considerando los temas, conceptos y variables relacionados con los intereses y objetivos del estudio, se identificaron varias categorías de análisis que dieron pauta a la revisión de la literatura, principalmente en el contexto universitario. Los resultados de esta etapa de trabajo dan sustento teórico a desarrollo del proyecto de investigación. La selección de materiales se determinó con base en las categorías, subcategorías e indicadores según el propósito del estudio y las tendencias emergentes en las áreas de interés, como se muestra en la Tabla 1.

Para el acceso de los documentos se realizaron búsquedas en repositorios de acceso abierto (ScienceDirect, SpringerLink, Redalyc, Dialnet, JEDM, Repositorio Nacional CONACYT, ERIC, DOAJ, SciELO, Google Académico y repositorios científicos de universidades).

Se realizó un filtrado de trabajos con base en la pertinencia de acuerdo con el tema de investigación tratado, acorde a las categorías identificadas. La muestra resultante quedó conformada por 58 obras: 7 libros, 45 artículos publicados en revistas indexadas, 2 trabajos presentados en eventos académicos y 4 tesis de grado, con fechas de publicación comprendidas entre 1997 y 2020.

Tabla 1 Sistema de categorías de análisis.

CATEGORÍA	SUBCATEGORÍA	INDICADOR
Rendimiento académico	Determinantes del rendimiento académico universitario Evaluación del rendimiento académico	Factores Indicadores de rendimiento
Educación mediada por tecnología	Tecnologías de la información y comunicaciones Ambientes virtuales de aprendizaje Diseño instruccional	Tipos de sistemas de información Características Beneficios Modelos
Minería de datos educativos	Estudios basados en minería de datos Predicción del rendimiento académico	Técnicas Algoritmos de clasificación y predicción Variables predictivas Validez
Intervención educativa	Modelos y estrategias de intervención educativa	Determinantes Estrategias Resultados

El desarrollo temático se presenta con base en las categorías de análisis definidas en la Tabla 1. Los ejes sobre los que gira el trabajo corresponden al rendimiento académico, sus determinantes y la forma en la que se puede evaluar; también el uso de la minería de datos educativos para realizar predicciones del rendimiento académico, por lo que se requiere conocer sobre técnicas específica y métodos para seleccionar variables y validar resultados. El análisis de intervenciones

educativas es necesario para conocer sobre modelos y estrategias empleadas previamente, y con base en sus resultados, se propongan nuevas alternativas de actuación. Finalmente, al conocer sobre educación mediada por tecnología se tiene el sustento para definir ambientes de aprendizaje adecuados a las intervenciones educativas, empleando para ello las tecnologías de la información y las comunicaciones más apropiadas.

El análisis de contenidos se efectuó tratando de identificar las características principales de los trabajos, propósitos, enfoques teóricos, metodologías, estrategias, técnicas, herramientas empleadas y resultados. A su vez, el material fue organizado de acuerdo con las categorías y, por cada una de ellas, se contrastaron los textos buscando puntos de encuentro y desencuentro, alcances, límites, tendencias y vacíos en los campos de conocimiento tratados por los investigadores.

El balance de lo encontrado tiene un carácter interpretativo y trata de dar sentido y coherencia a los trabajos previamente realizados, con énfasis en la detección de vacíos de saber en las obras desarrollados hasta el momento en el campo del conocimiento tratado y en consecuencia se justifiquen nuevas propuestas de investigación en esta línea temática. Es decir, la información derivada de este análisis tuvo el propósito de construir las bases para ampliar o complementar el área de conocimiento o generar temáticas emergentes de investigación.

## **2.2 Determinantes del Rendimiento Escolar**

En la obra de Lamas (2015) se presentó una revisión de trabajos relacionados con el rendimiento académico, desde una aproximación conceptual hasta los distintos y variados enfoques utilizados para entender el aprendizaje y los factores que lo afectan. Identifica algunos de los modelos empleados para analizarlo, por ejemplo, el modelo psicológico centrado en las características del estudiante, modelo integrador de teorías sobre el rendimiento académico, basado en la Teoría Cognitiva-Logro.

Por su parte, Moreno et al. (2015) analizaron la relación de la atención del docente, motivación y necesidades psicológicas básicas que determinan la percepción de competencia académica y las calificaciones de alumnos en una universidad de España, mediante un análisis de correlaciones bivariadas y de regresión estructural. Chong (2017) identificó y analizó los factores que inciden en el desempeño académico de los estudiantes de la Universidad Politécnica del Valle de Toluca (México), debido a que no se observaron mejoras en su rendimiento a pesar de contar con recursos tecnológicos y apoyos proporcionados por la institución; los resultados del análisis estadístico del cuestionario aplicado muestran que los estudiantes consideran la preparación de los profesores, sus ganas de superarse y, en menor medida, sus relaciones familiares, como los factores que contribuyen en su aprendizaje.

En la tesis de Dorio (2016), se desarrolló un estudio con enfoque mixto para determinar aspectos que determinan el rendimiento y la adaptación de los estudiantes a la universidad, utiliza categorías de variables relacionadas con la motivación por la carrera, la autoeficacia, el estrés académico, el apoyo del entorno, la satisfacción con las calificaciones recibidas y por la carrera. Se comprobó la influencia e interacción de factores personales e institucionales en los estudiantes de primer semestre, que determinan el proceso de transición y el rendimiento académico. Los estudiantes manifestaron la necesidad de contar con redes de apoyo sociales y académicas que les posibiliten el afrontar las dificultades de la adaptación a la escuela. Plantea la necesidad de elaborar mecanismos más flexibles y adaptables al contexto de las instituciones educativas (de universidades españolas) para prevenir y favorecer que los estudiantes que inician sus estudios universitarios puedan progresar y graduarse. Destaca que los programas de apoyo y orientación deben favorecer la integración y la adaptación al contexto universitario y a lo largo del proceso formativo.

La investigación realizada por Padua (2019), tuvo el propósito de identificar factores familiares e individuales relacionados con el bajo rendimiento académico, en una

universidad del norte de México. Se compararon grupos de estudiantes, 48 con promedio de calificaciones menores a 7.5, y 48 con promedio de calificaciones de 9.0 o mayor. Mediante cuestionarios tipo Likert y pruebas de correlación no paramétricas, se identificó influencia positiva del promedio con las variables maestría, trabajo, actitud ante el estudio, autoconcepto académico, creencias espirituales y estrategias de estudio; e influencia negativa con el control psicológico materno, imposición paterna y competitividad.

Los estudios analizados en esta dimensión muestran una gran diversidad de enfoques, determinados en gran medida por el campo de conocimiento desde donde se realiza el análisis. Sin embargo, hay coincidencias al considerar, por un lado, a las habilidades cognitivas, la motivación, la autoeficacia y la autorregulación del aprendizaje, además de las relaciones interpersonales y familiares, como factores determinantes del logro académico. Y, por otro lado, los elementos del entorno tienen un papel fundamental en el desempeño escolar, por ejemplo, las características institucionales y académicas de las escuelas. Una clave para lograr el éxito académico parece ser la búsqueda del equilibrio entre los diferentes factores que lo determinan.

### **2.3 Evaluación del Rendimiento Académico**

En la tesis de López-Ramírez (2015), se aplicó un método sistémico para evaluar el rendimiento académico, con resultados positivos en los casos de estudio planteados. Revisó alternativas comúnmente utilizadas en la evaluación del rendimiento desde la perspectiva de la calidad educativa. Retoma planteamientos formales de Ralph Tyler, considerado el precursor de la evaluación educativa, de sus trabajos desarrollados entre 1930 y 1945, en los que se considera a la evaluación como el proceso que determina el nivel de logros alcanzados y que se evidencia al observar cambios deseables en los comportamientos de los seres humanos.

Para York, Gibson y Rankin (2015), el éxito académico es uno de los constructos más ampliamente usados en investigación y evaluación educativa, en su investigación realizó una revisión del uso de estos conceptos en múltiples campos académicos. Encontraron que el logro académico, determinado por las calificaciones y el GPA (*Grade Point Average*, o promedio de calificaciones) son las medidas de éxito académico más ampliamente utilizadas; aunque en su modelo conceptual de éxito académico también considera el éxito profesional, la satisfacción, la persistencia, la adquisición de habilidades y competencias, así como el logro de resultados de aprendizaje; de igual forma proponen un modelo operacionalizado en donde presentan estas categorías con indicadores medibles en el contexto académico. Destacan la necesidad de considerar todos estos elementos, que determinan en su conjunto el éxito académico, para no limitar la calidad de las investigaciones y el entendimiento de lo que contribuye al éxito de los estudiantes.

Si bien varias de las investigaciones revisadas consideran las calificaciones como el principal elemento para evaluar el rendimiento académico (York et al., 2015), hay posturas con un enfoque más holístico en que se toman en cuenta los diferentes y complejos factores determinantes del rendimiento escolar a lo largo del todo el proceso formativo e inclusive en el ámbito profesional (López-Ramírez, 2015). Las diferentes posturas comúnmente dependerán del propósito con el que se realiza la evaluación.

#### **2.4 Tecnologías en el Contexto Educativo**

Diversos autores sitúan a la Tecnología Educativa (TE) como una disciplina que forma parte de la didáctica, constituida por elementos curriculares que favorecen el desarrollo de habilidades cognitivas en los sujetos, en contextos determinados, facilitando y estimulando la intervención mediada sobre la realidad, la captación y comprensión de la información por los alumnos y la creación de entornos diferenciados que propician el aprendizaje (Cabero, 2003; Torres y Cobo, 2017). La

TE implica tareas de utilización, organización, desarrollo, diseño y evaluación de las Tecnologías de la Información y Comunicaciones (TIC).

Las nuevas concepciones sobre la TE, sustentadas en la psicología cognitiva, asignan un nuevo rol a las TIC, como instrumentos potenciadores e impulsores de las habilidades cognitivas de los usuarios. De manera más reciente, la expansión y madurez de las nuevas tecnologías de las comunicaciones, las han convertido en herramientas que fomentan la interacción social y que su vez propician el aprendizaje colaborativo y cooperativo, lo que refleja las tendencias de los nuevos modelos educativos los cuales que sugieren su empleo en tres grandes áreas aplicativas: comunicación, conocimiento y participación (Cabero, 2015).

El abanico de opciones digitales que se han incorporado a los procesos de enseñanza y gestión educativa se ha incrementado sustancialmente como lo muestran varios reportes (Bates, 2015; Brooks, 2016; Sharples et al., 2014): web semántica, internet de las cosas, analíticas del aprendizaje, minería de datos educativos, realidad aumentada, computación en la nube, sistemas MOOC, gamificación, redes sociales, entre otras tecnologías que tienen un potencial enorme para mejorar los entornos y procesos de formación educativa.

Al respecto, Gros (2012) coincide en varios de los puntos mencionados, y, además, indica que es de suma importancia formular preguntas de investigación adecuadas para guiar en la dirección correcta los estudios. Para ello, recomienda repensar lo que se quiere saber de la investigación, pues es de mayor interés comprender el cómo se aprende mejor y qué saber si se aprende, con el uso de determinada innovación tecnológica; lo que llevará a entender cómo diseñar mejores contextos educativos para un aprendizaje más efectivo.

A pesar de existir gran cantidad de investigaciones en el campo de la TE, muchos de ellas carecen de fundamentación teórica y metodológica. El análisis de varios trabajos relacionados con TE, realizados por Bulfin et al. (como se citó en Cabero, 2016), revelan que la mayoría de las investigaciones utilizan métodos muy básicos

de trabajo y de análisis de la información, además de capacidades limitadas para la recopilación de datos cuantitativos. Investigaciones en TE precedentes han mostrado poca variedad en las metodologías empleadas, como se mencionó anteriormente, básicamente estudios descriptivos y comparativos.

El cambio de paradigma educativo centrado en las actividades que realiza el estudiante requiere replantear las investigaciones hacia un enfoque más holístico y sistémico, cuyo eje sea el diseño de los entornos de aprendizaje con elementos tecnológicos que mejoren el proceso de enseñanza-aprendizaje. Resulta entonces necesario combinar diferentes modalidades de estudios: cuantitativos y cualitativos, realizar estudios longitudinales, mejorar las técnicas de recogida de datos, emplear herramientas que mejoren la rapidez en el análisis de los datos, usar diversas variables que reflejen mejor los complejos entornos de aprendizaje reales.

De forma reciente, el término *e-research* se ha utilizado para agrupar a las nuevas metodologías de investigación derivadas del uso de Internet, sistemas de gestión de aprendizaje y el manejo de grandes volúmenes de datos digitales (*Big Data*), en este sentido Gros (2012), sugiere utilizar las analíticas del aprendizaje y la minería de datos para interpretar la gran cantidad de datos de los estudiantes generados en los nuevos entorno digitales, así como para registrar y analizar su progresión académica, predecir actuaciones futuras e identificar elementos problemáticos. El uso de estas técnicas y herramientas tecnológicas en TE, tiene implicaciones importantes pues, además de la aplicación de los métodos estadísticos tradicionales, se incluyen métodos deductivos y de aprendizaje automático en los que, a partir de los datos se descubren patrones, reglas e hipótesis. Es preciso entender las diferencias y características de los métodos estadísticos y los basados en minería de datos para orientar adecuadamente las investigaciones.

La incorporación de TE a la educación no debe ser un asunto de moda, su empleo debe estar justificado y soportado por las teorías del aprendizaje que sean acordes a la realidad del contexto educativo de los actores del aprendizaje. Debemos ser

cuidadosos al momento de plantear los objetivos de nuestros proyectos de investigación y seleccionar los métodos más adecuados para trabajar y validar los resultados. Cualquier metodología elegida puede ser de utilidad, siempre y cuando este bien justificada y se aplique correctamente. Nuevos enfoques de investigación pueden ser utilizados para comprender mejor lo que ocurre realmente en el proceso educativo de los estudiantes, generando la oportunidad de intervenir de manera oportuna en el diseño y reajuste del entorno de aprendizaje. Las nuevas metodologías para obtener información a partir de las interacciones de usuarios con aplicaciones y sistemas basados en Internet pueden ser de gran utilidad para identificar con rapidez los elementos que inciden en el aprendizaje, dar seguimiento al proceso, proporcionar recursos personalizados para los alumnos y facilitar la toma de decisiones estratégicas para definir mejores formas de actuación.

## **2.5 Minería de Datos Educativa y Rendimiento Académico**

A partir de la conformación de nuevas disciplinas y comunidades para el análisis de datos educativos, se han incrementado los trabajos relacionados con el estudio de las características de alumnos y sus ambientes escolares (Baker e Inventado, 2014). El uso de MDE y LA en el contexto educativo tiene el potencial de transformar los modelos existentes de enseñanza-aprendizaje al proporcionar nuevas herramientas de análisis, interacción e intervención (Aldowah et al., 2019). Asimismo, cuentan con métodos que pueden emplearse para el desarrollo de investigaciones aplicadas con el propósito de mejorar la calidad de la educación o de investigaciones puras que buscan mejorar el entendimiento del proceso de aprendizaje (Bakhshinategh et al., 2018).

Algunos de los trabajos analizados relacionados con el uso de minería de datos educativos y la predicción del rendimiento académico o problemáticas asociadas se describen a continuación: En su tesis doctoral, García (2015) desarrolló un modelo para determinar el rendimiento académico de alumnos universitarios en modalidad a distancia, en Ecuador; empleando técnicas multivariantes para la analítica del

aprendizaje (LA); también determinó el aporte de variables individuales, grupales y contextuales en el rendimiento escolar. Martínez et al. (2015) consideran al rendimiento académico como un factor crítico asociado a las altas tasas de deserción, analizaron resultados de bajo rendimiento académico en asignaturas del primer nivel en la Universidad Tecnológica Nacional, provincia del Chaco, Argentina; mediante técnicas de la DM determinaron perfiles y clasificaron alumnos con bueno y bajo rendimiento académico, con base en antecedentes de estudios previos y datos familiares.

Muñoz (2015) trabajó en su tesis doctoral aspectos de la mejora del rendimiento académico de estudiantes de enseñanza secundaria obligatoria en España, utilizó DM para el análisis de información tanto académica, personal y de pruebas psicométricas aplicadas a la población de estudio; siguió un enfoque Onto-Semiótico de la Didáctica, buscando mejorar las habilidades matemáticas y lingüísticas, que se consideran fundamentales para el desarrollo de otras asignaturas y de estudios posteriores a nivel bachillerato o universitarios. López et al. (2015) generaron un modelo para predecir el bajo rendimiento académico en alumnos de la Universidad Nacional de Colombia, usan datos de los cuatro primeros períodos de estudio y aplicaron los algoritmos Naive Bayes y Decision Tree para la tarea de clasificación.

También, Aldikanji y Ajami (2016) usaron MDE para descubrir correlaciones entre el promedio y algunos descriptores de estudiantes en la Universidad Virtual de Siria. Por su parte, Merchan y Duarte (2016) construyeron un modelo predictivo de desempeño académico usando datos geográficos y académicos de estudiantes del programa de Ingeniería en Sistemas de la Universidad del Bosque en Bogotá, Colombia; para aplicar varias técnicas de clasificación de la MDE (J48, Part y Ridor); las precisiones y reglas obtenidas son diferentes para cada algoritmo, aunque el clasificador estrato social es común; observaron que los resultados obtenidos están altamente relacionados con las características de la institución y su contexto. Menacho (2017) aplicó las técnicas de MDE de regresión logística, árboles de

decisión, redes bayesianas y redes neuronales, usando datos escolares de estudiantes matriculados en el curso de Estadística General de la Universidad Nacional Agraria La Molina (Lima, Perú), para predecir la clasificación final (desaprobado o aprobado); la red Naive de Bayes obtuvo la mejor precisión; sugiere considerar datos socio económicos para mejorar la precisión del modelo. Miranda y Guzmán (2017) determinaron las variables que llevan a un estudiante abandonar sus estudios universitarios, usando técnicas de MDE, los datos se obtuvieron de alumnos de las carreras de Ingeniería de la Universidad Católica del Norte en Antofagasta y Coquimbo (Chile); encontraron que las variables que mejor explican la deserción de un estudiante son: las razones socioeconómicas y el puntaje de ingreso a la universidad; para el modelado de los factores de deserción utilizaron red bayesiana, árbol de decisión y red neuronal.

Kumar, Singh y Handa (2017) hicieron una revisión de trabajos realizados de 2007 a 2016, para identificar las principales técnicas de minería de datos utilizadas para predecir el progreso y desempeño de estudiantes e identificar los atributos más importantes para las predicciones. La mayoría de las investigaciones han utilizado los algoritmos de Decision Tree, Naive Bayes, Rule-Based, Artificial Neural Networks y K-Nearest Neighbor; y las mejores predicciones se han logrado con una combinación de atributos personales, socio-económicos y de estudios previos. Posteriormente, Kumar y Singh (2017) evaluaron varias técnicas de minería de datos educativas, con el propósito de determinar las mejores. Utilizaron datos tanto académicos como familiares, de alumnos de la Universidad Himachal Pradesh, India. Concluyeron que el algoritmo de modelado predictivo Random Forest da mejores resultados que el Decision Tree, Naive Bayes, Bayes Network y árboles de clasificación y regresión. También emplearon técnicas de validación cruzada y conjuntos de entrenamiento.

Costa et al. (2017), dada la creciente preocupación por el bajo rendimiento de los alumnos en los primeros cursos de nivel universitario, realizaron un estudio de efectividad de las técnicas de minería de datos (Decision Tree, Support Vector

Machine, Neural Network y Naive Bayes), para la predicción oportuna del posible fracaso escolar en cursos de programación en la Universidad Federal de Alagoas (UFAL), Brasil; concluyeron que las técnicas analizadas son suficientemente efectivas para el propósito planteado. Rico y Sánchez (2018) diseñaron y automatizaron un modelo predictivo del rendimiento académico de estudiantes del Instituto Politécnico Nacional (IPN), en México. El modelo utilizó las calificaciones de cinco actividades académicas y la calificación final; también se construyó una plataforma que facilita la implementación del modelo para predecir automáticamente el desempeño académico de nuevos estudiantes, para ello se programó el algoritmo de Naive Bayes. Hamoud, Hashim y Awadh (2018) construyeron árboles de clasificación (J48, Random Tree y RepTree), para descubrir patrones de desempeño académico de estudiantes universitarios (Iraq), utilizaron un cuestionario sobre aspectos sociales, de salud y académicos.

El uso de algoritmos de MDE fue utilizado en el trabajo de Aziz, Hafieza y Ahmad (2014) para determinar el desempeño académico (pobre, promedio, bueno) de alumnos de primer semestre de una carrera en ciencias de la computación en Malasia. Los métodos para generar los modelos predictivos fueron Naive Bayes, Rule Based, y Decision Tree. Utilizaron 5 variables relacionadas con aspectos socioeconómicos y como variable dependiente el promedio de calificaciones en primer semestre (GPA, *Grade Point Average*). En las pruebas de validación emplearon variantes con diferentes porcentajes de valores de entrenamiento y de validaciones cruzadas, con un 68.8% de exactitud como valor más alto.

Por su parte, Miguéis et al. (2018) aplicaron técnicas de MDE para diseñar modelos predictivos de clasificación temprana de estudiantes universitarios según su potencial rendimiento académico. Los resultados empíricos indicaron que el algoritmo Random Forest fue mejor que las otras técnicas de clasificación utilizadas, al obtener un 95% de exactitud en las predicciones. Sin embargo, los algoritmos Decision Trees, Support Vector Machines, Bagged Trees y Boosted Trees, obtuvieron altos valores de exactitud, por lo que los modelos predictivos obtenidos

también son efectivos para su aplicación en el contexto académico, Naive Bayes tuvo el peor desempeño. Los atributos con mayores pesos fueron el promedio de calificaciones de los exámenes de ingreso y los promedios de calificaciones al primer año, con lo cual se generaron modelos confiables para predecir el rendimiento académico para el resto de la carrera. En menor medida tuvieron influencia la cantidad de créditos obtenidos en el primer año y el promedio de calificaciones en el nivel educativo previo, así como otros factores socioeconómicos.

Para la creación de un sistema de detección temprana de alumnos en riesgo, Berens et al. (2019), utilizaron el meta-algoritmo AdaBoost para combinar análisis de regresión, redes neuronales y árboles de decisión en la predicción de abandono escolar de alumnos de universidades de Alemania. La precisión de la predicción, usando datos demográficos y calificaciones disponibles al final del primer semestre, fue del 67%, y aumento a 80% en el cuarto semestre. Los resultados permiten identificar estudiantes en riesgo y ubicarlos en programas de apoyo en las universidades, además de servir de punto de referencia para probar la efectividad de las intervenciones educativas.

A pesar de las grandes posibilidades de la MDE y las AA, lo cierto es que en el contexto de la educación superior su empleo aún es incipiente, como se menciona en Aldowah et al. (2019). Es necesario realizar más estudios y desarrollar proyectos de mejora educativa en donde se aproveche mejor su potencial, tratando de analizar de forma más fina los procesos educativos (Villanueva et al., 2018). El uso de estas herramientas tecnológicas sigue siendo un reto en el contexto universitario para mejorar la calidad de la toma de decisiones de instructores y administrativos, así como para identificar y analizar muchos de los elementos del entorno académico de manera eficiente y rápida.

En relación con los trabajos que consideran la integración de modelos predictivos en sistemas informáticos, Marques, Hobbs y Graf (2014) destacan las dificultades para obtener información sobre el comportamiento y las actividades de los

estudiantes en sistemas de gestión de aprendizaje (LMS, *Learning Management Systems*), por lo que propusieron un modelo de predicción de riesgo académico que considera el porcentaje de actividades entregadas, promedio de calificaciones, porcentaje de recursos vistos y asistencia, durante el desarrollo de un curso. El modelo fue generado mediante el algoritmo de árbol de decisión y en la implementación se verifica cada factor de riesgo en la secuencia de importancia indicada por el árbol. Con base en el modelo, se desarrolló una herramienta, en el *LMS Moodle*, para generar reportes de estado de riesgo. Los resultados indicaron que un 85% de estudiantes fueron correctamente clasificados.

Por su parte, Rico y Sánchez (2018) diseñaron y automatizaron un modelo predictivo del rendimiento académico de estudiantes del Instituto Politécnico Nacional (IPN), en México. El modelo utilizó las calificaciones de cinco actividades académicas y la calificación final; también se construyó una plataforma que facilita la implementación del modelo para predecir automáticamente el desempeño académico de nuevos estudiantes, para ello se programó el algoritmo de Naive Bayes.

Los resultados derivados de muchos de los estudios que emplean la MDE se limitan a la generación de modelos y la descripción de reglas que caracterizan los factores prominentes de desempeño escolar; otros más, se orientan a la comparación del rendimiento o precisión de las diversas técnicas y algoritmos que conforman el repertorio de herramientas de la MDE. De acuerdo con los trabajos revisados, se observa que las técnicas de MDE más usadas, en la predicción del desempeño, caracterización de los estudiantes, bajo desempeño y deserción, se ubican dentro de los métodos de regresión y clasificación; pero también se han utilizado métodos de agrupamiento y reglas de asociación (Alyahyan y Düşteğör, 2020; Bakhshinategh et al., 2018).

Desde nuestro punto de vista, ha quedado pendiente la implementación de los modelos, mediante sistemas o programas computacionales, para automatizar y

facilitar el análisis de nuevos casos, de manera que se consolide como una herramienta de uso regular en la toma de decisiones académicas y de apoyo para la generación y aplicación temprana de estrategias de intervención en situaciones de riesgo escolar. En los trabajos en donde si se implementan sistemas, se consideran sólo datos generados durante el desarrollo de un curso, por lo que se deben evaluar algunas actividades antes de poder aplicar los modelos predictivos.

## **2.6 Estrategias de Intervención Educativa**

En palabras de Barraza (2010, p. 24), “la Propuesta de Intervención Educativa es una estrategia de planeación y actuación profesional que permite a los agentes educativos tomar el control de su propia práctica profesional mediante un proceso de indagación-solución”.

En el ámbito de la intervención educativa, Masino y Niño (2016) hicieron una revisión de intervenciones para la mejora de la calidad de la educación y el aprendizaje de estudiantes en países en desarrollo. Explican que, de acuerdo a la tipología de la teoría del cambio, existen tres elementos primordiales para impulsar el cambio en la calidad de la educación: (1) intervenciones del lado de la oferta, suministrando los recursos físicos y humanos, y materiales de aprendizaje; (2) políticas de incentivos, para influenciar preferencias y comportamientos de maestros, hogares y estudiantes; (3) intervenciones de gestión comunitaria, implementadas a través de reformas de descentralización, difusión del conocimiento y una mayor participación de la comunidad en las decisiones del sistema educativo. Los resultados generales indican que las intervenciones, para mejorar el rendimiento y el aprendizaje de los estudiantes, son más efectivas cuando se emplean dos o más elementos impulsores del cambio para diseñar políticas educativas.

Albarracín y Montoya (2016) analizaron investigaciones relacionadas con programas de intervención para estudiantes universitarios con bajo desempeño académico. Los programas y sus resultados se ubicaron en tres categorías,

definidas por las 1) condiciones personales, 2) condiciones sociales, y 3) condiciones institucionales. La mayoría de los programas han utilizado estrategias de autorregulación, tutorías y estrategias para desarrollar áreas específicas. Se encontraron pocos programas para intervenir en los aspectos socioeconómicos, familiares y sociales de los estudiantes, a pesar de que se reconoce su importancia. En cuanto a las condiciones institucionales, recomiendan continuar con la cualificación de los docentes y utilizar recursos para mejorar los métodos y ambientes en los que se desarrolla la enseñanza y el aprendizaje. Las intervenciones que incluyeron Tecnologías de la Información (TIC), lograron un mayor impacto y efectividad, además de una serie de beneficios como la eliminación de barreras espacio-temporales, variedad de modalidades de comunicación y una mayor flexibilidad y motivación.

Un caso de utilización de TIC se describe en la investigación realizada por Benítez, Barajas y Hernández (2014), en donde se aplicó una estrategia instruccional integradora para la comprensión de la lectura, utilizando un entorno virtual para desarrollar habilidades de comprensión de textos en Ciencias Sociales y Administración en los estudiantes del segundo semestre de las licenciaturas de Contaduría Pública y Administración de una Institución de Educación Superior de San Luis Potosí, México. Los resultados del experimento realizado mostraron que las estrategias de intervención mediante el uso del entorno virtual de aprendizaje contribuyeron a mejorar las habilidades de comprensión de textos.

La investigación de Silva (2011) pone de relieve la importancia del primer año universitario para realizar investigaciones educativas que den a conocer los perfiles de los estudiantes y se diseñen programas de apoyo para mejorar la experiencia en las aulas. Considera que la intervención en el primer año de estudios es fundamental para reducir problemáticas de reprobación, rezago y deserción, y favorecer trayectorias escolares exitosas. Proponen los seminarios de primer año (*freshman seminars* o *first-year-seminar*), la cual ha sido una alternativa exitosa en escuelas de Estados Unidos para brindar apoyo y facilitar la transición. Estos

seminarios están incluidos en el currículo y buscan principalmente la orientación, además de desarrollar habilidades académicas y proporcionar información sobre recursos y servicios de la universidad.

Más específicamente, los contenidos pueden ser sobre habilidades de estudio, recursos y servicios, manejo de tiempo, planificación académica y desarrollo de pensamiento crítico (Stuart y Linder, como se cita en Silva, 2011)). La estrategia no es limitativa, y se invita a emprender procesos de innovación para la transformación pedagógica efectiva.

Velázquez y Rodríguez (2014), introducen el concepto de tutoría de asignatura, la cual es una modalidad de intervención tutorial en la que se planifican acciones de formación académica y profesional de manera proactiva, a diferencia de la tradicional tutoría de acompañamiento en la que el profesor reacciona a las problemáticas académicas o personales que detecta en los estudiantes. La intervención consistió en aplicar entrevistas diagnósticas para conocer las expectativas no satisfechas de formación profesional de los alumnos. El profesor-tutor dirigió a los alumnos en el diseño de una solución a una problemática de desarrollo de software detectada y revisaron temas no cubiertos en el currículo de la licenciatura, como inteligencia de negocios, cuadro de mando integral y desarrollo ágil de software. Adicionalmente construyeron un sistema para generar un indicador sobre la práctica docente, que les fue de utilidad para detectar situaciones que pudieran afectar el desarrollo de las clases.

En el área de la salud, Guerra y Borrillo (2018) recopilaron y sistematizaron estudios relacionados con la influencia de la tutoría en el rendimiento académico de estudiantes de Ciencias de la Salud, en Sevilla, España. Concluyeron que las tutorías son eficaces si el ambiente en que se imparten es adecuado y los tutores están bien preparados y con la dedicación suficiente. Burgos et al. (2018), después de generar un modelo predictivo de rendimiento académico a partir de datos de actividades de un curso en línea, en la Universidad Abierta de Madrid, España;

diseñaron un plan de tutoría para reducir la deserción, los resultados mostraron una reducción significativa de la deserción en los cursos analizados, en comparación con cursos de años previos en donde no se habían implementado los mecanismos de predicción e intervención.

La revisión de trabajos realizada por Zeng, Ju y Hord (2018), tuvo el propósito de identificar los tipos de intervenciones que ha sido usadas para apoyar a estudiantes universitarios con dificultades de aprendizaje entre 2000 y 2016. Definieron cuatro categorías de tipos de intervenciones utilizadas para mejorar el éxito académico: uso de tecnología de asistencia (16.7%), asistencia directa (tutorías o asesorías) (8.3%), instrucción estratégica (50%) y programas integrales de apoyo (25%). La asistencia con tecnología se empleó para compensar dificultades de lectura y escritura mediante dispositivos o sistemas que fueron adaptados o personalizados según las necesidades individuales. Los estudios donde se utilizó la instrucción o enseñanza estratégica en las áreas de escritura, lectura, evaluación y dominio de la información, mostraron efectividad para mejorar sus procesos cognitivos, como la autoeficacia, la autorregulación y el conocimiento metacognitivo. Los programas integrales de apoyo estuvieron enfocados en promover habilidades académicas y la autodeterminación.

La propuesta de intervención en la iniciativa que presentaron Jayaprakash et al. (2014), consistió en el envío de mensajes de alerta a estudiantes en riesgo de varias escuelas de educación superior en Estados Unidos. De acuerdo con los autores, la simple notificación de un posible riesgo académico puede afectar significativamente el comportamiento del estudiante y hacer que busquen ayuda. Su propuesta sigue el principio básico de establecer comunicación continua entre los estudiantes y la facultad. La detección o clasificación de alumnos en riesgo se logró mediante modelos generados con minería de datos; y las intervenciones se lograron con la implementación de dos estrategias: envío de mensajes de riesgo académico exclusivamente y envío de mensajes acompañado del uso de un sistema de apoyo académico en línea.

## **2.7 Líneas de Investigación Emergentes**

En la revisión del *corpus* de la literatura observamos un gran interés por entender los determinantes del rendimiento académico y para ello se han empleado varias metodologías, técnicas y herramientas de análisis. Así mismo, desde el punto de vista psicopedagógico, la inteligencia, las aptitudes mentales o habilidades cognitivas y características de la personalidad, se han considerado como relevantes en los estudios sobre rendimiento escolar.

Dado que el rendimiento académico está asociado con la calidad de la educación, reflejada en sus diversos indicadores y también es el resultado de la relación entre factores personales, sociales, institucionales o del entorno, en varias de las obras revisadas se recomienda realizar estudios para diagnosticar la situación particular de los estudiantes, haciendo énfasis en que cada institución tiene sus propias características y que para ciertos contextos los resultados y modelos obtenidos en otros estudios no son generalizables. En varios de los trabajos se han usado los métodos de análisis estadístico para verificar la relevancia de las variables involucradas en el rendimiento escolar. Las diversas investigaciones dan cuenta de los múltiples factores y variables que se deben considerar, para obtener esquemas conceptuales más generales.

Se vislumbra un gran espacio de oportunidad para aplicar nuevas metodologías de investigación, soportadas por las tecnologías de la información y las comunicaciones, particularmente aquellas que hacen uso de la minería de datos y las analíticas del aprendizaje, pues sus técnicas pueden acelerar los procesos de generación de conocimiento a partir de los datos digitales producidos en masa por los sistemas y ambientes virtuales de aprendizaje. Otra área de interés, para el desarrollo de nuevas investigaciones, puede ir dirigida a analizar como el reforzamiento de determinados factores pueden compensar la carencia o debilidad en otros aspectos que inciden en el rendimiento académico.

En relación con la evaluación del rendimiento académico, la mayoría de los trabajos en investigación educativa han considerado las calificaciones de las materias o el promedio de calificaciones como el principal indicador de rendimiento o éxito, sin embargo, es recomendable considerar otras categorías de indicadores para lograr una evaluación más apropiada, pues la naturaleza misma del rendimiento académico es compleja y requiere ser analizada desde diferentes perspectivas. Considerando el alcance y objetivos de la investigación, enfocada en determinar el rendimiento académico de alumnos de primer grado, en una primera instancia, se considera conveniente utilizar indicadores de rendimiento inmediato, como son las calificaciones o la cantidad de materias reprobadas; en una etapa posterior, se considerarán categorías e indicadores relacionadas con la adquisición de habilidades y competencias académicas. Finalmente, se tiene contemplada la implementación de estrategias de intervención basadas en tecnologías educativas, en donde el uso de la minería de datos educativos y las analíticas del aprendizaje pueden convertirse en herramientas decisivas para dar seguimiento a las actividades de los cursos o programas de tutoría y asesoría, a la vez que se constituyen como una excelente alternativa para proporcionar retroalimentación en tiempo real a todos los actores involucrados en el proceso educativo.

En el ámbito de la tecnología en el contexto educativo, la mayoría de las universidades cuentan con sistemas de información para la gestión de datos escolares de los alumnos, o con plataformas para la gestión del aprendizaje que usan para la impartición de cursos presenciales con soporte en línea, y algunos cursos en modalidad virtual; además, existen diversos sistemas para la evaluación docente y la evaluación curricular, entre otros, que generan de manera rutinaria gran cantidad de datos almacenados en formato digital. Todo este caudal de información, desafortunadamente, no se ha empleado consistentemente para realizar investigación educativa u obtener información novedosa y relevante sobre los alumnos y sus procesos formativos, de manera que se pudieran detectar situaciones problemáticas e intervenir para mejorar la calidad de la educación.

Adicionalmente, como se destacó en la literatura revisada sobre el tema, existen grandes avances en las tecnologías de la información y comunicaciones, con las capacidades para soportar la implementación de novedosos diseños instruccionales y didácticos, que, sin embargo, se han visto detenidos por el desconocimiento o por la falta de pericia de los docentes al manejar estas tecnologías. El reto actual es superar estas limitantes e integrar de manera justificada y armoniosa estos elementos de innovación tecnológica en el desarrollo de cursos, soportados por sistemas de gestión de aprendizaje y diseñados con nuevos enfoques didácticos, como el *e-learning* o el *b-learning*. Por ejemplo, se pueden diseñar cursos basados en estas modalidades educativas para adaptarse a las nuevas necesidades del complejo entorno social y educativo de nuestros días.

La Minería de Datos Educativos, ha sido empleada en los últimos años en un número creciente de investigaciones. Sus métodos para explorar, visualizar y analizar de manera más ágil los datos de los entornos educativos, han permitido entender desde otras perspectivas los diversos escenarios de las instituciones educativas, los procesos de enseñanza-aprendizaje, el desempeño de los estudiantes y las condiciones en las cuales ellos aprenden. Dada la complejidad y multidimensionalidad del rendimiento académico, en el contexto de universitario, es improbable llevar a cabo estudios que mapeen cada una de las causas que influyen en el mismo, a pesar de ello, desde nuestro punto de vista, es factible encontrar algunos factores relevantes tomando como insumo la información digital disponible sobre el historial académico y los datos extraescolares de los alumnos, de manera que se generen modelos con un aceptable nivel de precisión, lo que permitirá atisbar el futuro académico de los estudiantes y con ello, en la medida de lo posible, atender algunas de las más importantes causas del bajo rendimiento escolar en su contexto educativo particular.

En su mayoría, los estudios que emplean minería de datos educativos, aportan poca información relativa a procesos de intervención una vez determinados los modelos o reglas que establecen las relaciones entre factores de posibles problemáticas

académicas mediante la MDE. En este sentido, es necesario transitar hacia un enfoque más integral, de manera que, a partir de los modelos predictivos, contruidos con información académica y no académica, disponibles desde el momento mismo de ingreso a la institución, se puedan identificar a los alumnos en riesgo y con ello, actuar con mayor oportunidad, aplicando estrategia de intervención educativa diseñadas a la medida de las necesidades.

De acuerdo con las obras analizadas, se ha observado que las intervenciones educativas implementadas suelen ser más efectiva cuando se emplean diversas estrategias de manera simultánea para mejorar el aprendizaje de los estudiantes. De igual forma, ha habido resultados exitosos cuando las medidas se han aplicadas de forma planificada y cuidadosa, además de ser necesario el contar con información oportuna y precisa sobre los perfiles educativos que se desean desarrollar, para diseñar las intervenciones con base en los conocimientos y habilidades requeridas por la población estudiantil. La aplicación de estrategias no debe ser limitativa, es recomendable emplear todas aquellas que sean factibles de implementar para lograr los propósitos perseguidos.

Dichas estrategias podrán ser definidas a nivel de aula o de escuela y requerirán la coordinación de actividades con directivos, responsables de programas y profesores-tutores para su diseño y aplicación. El empleo de la tecnología educativa es deseable en las intervenciones educativas, dado su potencial y tomando en cuenta los resultados positivos de diversos estudios. Además, en muchas circunstancias, como en el caso de la educación a distancia, es la alternativa más viable para poder incidir de manera significativa en el aprendizaje y el rendimiento académico, por tal motivo debe ser considerada al planificar cualquier estrategia de intervención educativa.

Dada la compleja naturaleza del constructo rendimiento académico, se requiere de una metodología holística e integradora, que contemple todas las etapas del proceso educativo y los elementos más pertinentes para mejorar los resultados

académicos, y coadyuvar a un tránsito escolar con menores contratiempos. Con el desarrollo del proyecto, se espera que se reconozcan las bondades del empleo de nuevas metodologías y herramientas tecnológicas en el ámbito académico para analizar datos y obtener información de forma más eficiente y oportuna; y con ello se dé pie a su adopción en otros procesos educativos y en otros espacios, promoviendo con ello la innovación y la mejora de la gestión educativa en la Institución.

### **3. MARCO CONCEPTUAL Y TEÓRICO**

En este capítulo se presentan los principales referentes conceptuales y teóricos relacionados con la investigación. Se parte de una descripción de la perspectiva teórica del aprendizaje desde la que se aborda el estudio, para explicar luego varios aspectos relacionados con el rendimiento académico, considerando la transición universitaria y sus problemáticas asociadas como temas de interés, ante lo cual es ineludible el comentar sobre los procesos de intervención educativa. Posteriormente, se describe como la educación mediada por tecnología es importante en nuestros días y cómo debe ser llevado a cabo el diseño instruccional de cursos para aprovechar los recursos tecnológicos. Finalmente, se presenta un panorama general de la minería de datos educativos como herramienta que puede facilitar el análisis de datos y la generación de modelos predictivos del desempeño académico.

#### **3.1 Perspectiva del Desarrollo del Estudio**

El aprendizaje ha sido abordado desde múltiples perspectivas y se han propuesto varias teorías para explicar cómo se aprende y qué factores intervienen en el proceso. Desde un enfoque cognoscitivo “el aprendizaje es un cambio perdurable en la conducta o en la capacidad de comportarse de cierta manera, el cual es resultado de la práctica o de otras formas de experiencia” (Schunk, 2012, p. 3).

La Teoría Cognoscitiva Social (TCS) contempla tres elementos principales como agentes que moldean el aprendizaje y las conductas; al respecto, Bandura (como se citó en Schunk, 2012) propuso el modelo de causalidad de reciprocidad triádica, para analizar la conducta humana como un marco teórico en donde se dan interacciones recíprocas entre conductas, variables ambientales y factores personales, como las cogniciones. De acuerdo con esta postura, factores de comportamiento, factores ambientales (extrínsecos) y factores personales (intrínsecos) interactúan de forma dinámica para producir o afectar el aprendizaje. Asimismo, la TCS determina que las nuevas experiencias serán evaluadas por el

aprendiz mediante el análisis de sus experiencias pasadas. Por lo tanto, el aprendizaje constituye el resultado de una evaluación exhaustiva de la experiencia actual en comparación con el pasado.

### **3.2 Rendimiento Académico**

El rendimiento académico también denominado aptitud escolar, desempeño académico o rendimiento escolar implica el logro de objetivos o metas de aprendizaje establecidas en programas de estudio. Puede ser entendido como un nivel de conocimientos demostrado, por lo que implica procesos de evaluación. El rendimiento varía de acuerdo con las circunstancias, condiciones orgánicas y ambientales que determinan las aptitudes y experiencias. Cuando se produce un desfase entre el rendimiento académico obtenido y el rendimiento que se espera del alumno, se habla de rendimiento discrepante o bajo rendimiento académico (Lamas, 2015).

El rendimiento académico es la expresión de las capacidades y las características psicológicas del alumno desarrolladas y actualizadas a través del proceso de enseñanza aprendizaje que le posibilita obtener un nivel de funcionamiento y logros académicos a lo largo de un período o semestre, que se sintetiza en un calificativo final (cuantitativo en la mayoría de los casos) evaluador del nivel alcanzado (Chadwick, 1979, como se citó en Molina, 2015).

Dado el carácter multicausal del rendimiento escolar, hay una gran diversidad de factores personales, sociales, e institucionales en variedad de contextos geográficos y temporales que intervienen y explican el proceso de aprendizaje y sus resultados. Identificar el tipo de influencia de los factores y las relaciones entre las distintas categorías de variables, es clave en el desarrollo de investigaciones con este objeto de estudio (Muñoz, 2015).

En este sentido, se han realizado numerosos estudios para entender los factores del rendimiento escolar desde diversas áreas del conocimiento como la psicología,

la sociología y la pedagogía; y, recientemente, se han empleado técnicas de la minería de datos educativos.

### **3.2.1 Influencia de Variables Socioemocionales**

La variable rendimiento académico puede ser explicada por una multitud de causas tanto psicológicas como sociológicas, por ejemplo, inteligencia, sexo, aptitud, antecedentes académicos, antecedentes familiares, características de la escuela, resultados de evaluaciones previas, métodos de enseñanza, entre otros (Kerlinger y Lee, 2002).

En años recientes, las instituciones educativas han visto la importancia de las habilidades socioemocionales (HSE) como determinantes del desempeño académico y del logro de metas sociales, de tal suerte que se han modificado los instrumentos de evaluación de conocimientos, competencia y habilidades que se aplican a los egresados de nivel medio superior y superior, además de que se avanza en propuestas para incluirlas de manera explícita en los diseños curriculares (CENEVAL, 2021; Pía Otero y Pais, 2017).

El modelo de habilidades socioemocionales de Farrington et al. (2012), contempla cinco grandes factores para clasificar las habilidades socioemocionales. Dicho modelo propone una relación en la que a mayor desarrollo de las HSE se potencia el desempeño académico, así como una mayor perseverancia y continuidad de los estudios.

Como se observa en CENEVAL (2021), las principales HSE incluidas en el modelo e incorporadas al Examen General de Egreso de Licenciatura (EGEL Plus) son:

- Gusto por la escuela: pensamientos académicos relacionados con el disfrute que tienen los estudiantes sobre su escuela y las actividades que realizan en ella.

- Cooperación (habilidades sociales): trabajar en equipo para lograr una tarea común con un conjunto de personas.
- Perseverancia académica (persistencia): tendencia a terminar los trabajos académicos a pesar de los obstáculos o distractores.
- Metacognición: estrategias de aprendizaje, procesos y tácticas que se utilizan para recordar información; estrategias para monitorear el propio aprendizaje y pensamiento, así como para autocorregirse.
- Compromiso académico: conductas académicas relacionadas con actividades que se asocian a ser un buen estudiante, tales como asistir a clase, hacer tareas, organizar materiales, participar en clase y estudiar.

### ***3.2.2 Medición del Rendimiento Académico***

Desde una perspectiva holística, el éxito académico en la universidad es determinado por la preparación de sus estudiantes, el tiempo de egreso, la inserción laboral, los programas de investigación, la proyección social, entre otros, por lo que la evaluación del rendimiento académico considera la suma de los diferentes y complejos factores que intervienen en el estudiante. Por ejemplo, factores personales, sociales y propios de las instituciones de educación superior que forman el conjunto de elementos determinantes del rendimiento escolar (López-Ramírez, 2015).

Para García (2015), existe una distinción entre el rendimiento inmediato, en el que se consideran las notas o calificaciones de las materias; y el rendimiento mediato, el cual incluye logros personales y profesionales. También destaca que es frecuente el uso de los siguientes indicadores para determinar el nivel de rendimiento académico: calificaciones, pruebas objetivas, cantidad de materias aprobadas y cantidad de créditos acumulados.

En el plano del rendimiento mediato York et al. (2015), además del logro de resultados de aprendizaje, consideran en su modelo conceptual de éxito académico

otros elementos como el éxito profesional, la satisfacción, la persistencia, la adquisición de habilidades y competencias.

El desarrollo de habilidades y competencias académicas resulta crucial en los primeros años de estudio. Las habilidades metacognitivas dotan a las personas de la capacidad de reflexionar sobre lo que necesitan para aprender y sus logros cognitivos; también, les permite articular lo aprendido con experiencias y conocimientos previos que les ayuda en la solución de nuevas problemáticas (Chaves et al., 2015). La habilidad de autodirección del aprendizaje, por su parte, es fundamental en la educación ya que permite a los aprendices regular la tarea por realizar, definir metas y planes de actuación, así como tácticas de estudio, y adaptarse antes situaciones cambiantes o adversas (Schunk, 2012).

El promedio de calificaciones es un importante predictor para saber si un estudiante concluirá sus estudios, ya que, además de aspectos cognitivos, normalmente toma en cuenta otros factores no cognitivos. Es necesario identificar e incluir en el cálculo de las calificaciones, a aquellos atributos que sean determinantes en el desempeño académico. Por ejemplo, actitudes, hábitos de estudio, atención, manejo de tiempo, estrategias metacognitivas y sociales, que les permitan resolver problemas académicos y de su entorno (Farrington et al., 2012).

La evaluación del aprendizaje es concebida de diferentes formas por los actores de los procesos educativos. En los hallazgos de la investigación realizada por Hidalgo y Murillo (2017), se manifiestan cuatro posturas sobre la evaluación: como un proceso irrelevante, como una oportunidad de mejora, como herramienta de rendición de cuentas de la escuela y como un instrumento de rendición de cuentas del propio estudiante. Independientemente de la concepción adoptada, la evaluación requiere de la operacionalización de las diferentes variables implicadas en el proceso evaluativo para obtener una medida del éxito académico logrado por los estudiantes.

### **3.2.3 Bajo Rendimiento y Riesgo Académico**

Para aclarar algunos conceptos que pudieran ser confusos Carpio et al. (2018) precisa algunas definiciones relacionadas con el rendimiento y el riesgo académico:

El riesgo académico se define como un estado caracterizado por la concurrencia de factores que tornan al estudiante propenso a colocarse en alguna de las condiciones siguientes:

- Rezago escolar: entendido como el desfase temporal entre los logros planeados en los tiempos curriculares y los alcanzados por los estudiantes en un momento determinado.
- Bajo rendimiento escolar: referido a la mala calidad de los logros de los estudiantes.
- Bajo nivel de aprovechamiento escolar: se refiere a un aprendizaje insuficiente para cubrir los objetivos programados en las asignaturas.
- Fracaso escolar: es el incumplimiento irreversible de las metas escolares y de aprendizaje, imposibilita la continuación a otro nivel educativo.

Para estimar el rendimiento escolar se emplean comúnmente los promedios de las calificaciones obtenidas en las distintas asignaturas cursadas, o mediante la razón asignaturas cursadas / asignaturas aprobadas. Algunos de los estudios para determinar predictores de riesgo académico, como el de Baker et al. (2015), han considerado como alumno en riesgo académico a aquellos con promedios de calificaciones iguales o menores a un 73%. En el mismo sentido, en la investigación de Padua (2019), se trabajó con alumnos universitarios con bajo promedio de calificaciones y para ello se consideró un puntaje menor a 7.5 como indicador de tal estado académico. En ambos casos, una vez identificados los alumnos en riesgo, se les brindaba apoyo, tratando de indagar en factores familiares e individuales para atender las causas y reducir el riesgo académico.

### **3.3 Transición Universitaria**

La entrada a la universidad representa el inicio de una nueva etapa en la vida de la persona. Tiene múltiples significados según el punto de vista del estudiante, considerando sus características, su formación social y cultural, sus intereses, sus capacidades, su desarrollo, etc. y la institución que lo recibe, según su tamaño, tipo de organización, estructura, etc. La transición universitaria es un constructo complejo y multifactorial con factores entrelazados de tipo personal, institucional y contextual. También es dinámica y no lineal, susceptible a las vivencias que experimenta el estudiante como persona y debido a las condiciones sociales, políticas y económicas del contexto, por lo que requiere una visualización holística en su entendimiento (Dorio, 2016).

El mismo autor señala que el primer año de universidad es el momento más significativo de la experiencia universitaria pues representa el inicio de un período de cambio clave en la vida del estudiante, tanto personal como académico. Se destaca que los abandonos de los estudiantes se concentran en mayor medida en este período inicial ante la dificultad para adaptarse a la nueva situación. Sin embargo, también se señala que las experiencias académicas y sociales positivas iniciales refuerzan y motivan la acción de continuar y persistir.

Por ello, una responsabilidad institucional es comprender las necesidades de los estudiantes que llegan a la universidad, de igual forma, dar el seguimiento adecuado para saber cómo transcurre su integración y adaptación al contexto universitario, y sobre todo identificar aquellos factores que facilitan o inhiben el proceso de transición. La implementación de mecanismos que recuperen dicha información contribuirá al diseño de programas de apoyo a estudiantes de primer curso, que procuren establecer bases sólidas para el desarrollo de trayectorias escolares exitosas.

### **3.4 Intervención Educativa**

La intervención educativa está asociada con la investigación y reflexión de los procesos del ejercicio docente. Busca sistematizar las acciones, relaciones y significaciones de la práctica educativa, y además de diagnosticar problemáticas y explicar sus causas, también propone alternativas de transformación y cambio en el entorno y las prácticas educativas. Una de las metodologías más utilizadas en investigación educativa es la investigación-acción, la cual consiste en el estudio de una situación social con el fin de mejorar la calidad de la acción dentro de la misma. Se ubica en el paradigma cualitativo o interpretativo, y va más allá del enfoque de investigación de cuantificar y controlar, ya que durante el proceso establece una relación con lo que es investigado, a través del acto comunicativo y ético (Álvarez y Álvarez, 2014).

Una propuesta de intervención puede ser de Actuación docente, si el profesor es el principal diseñador e implementador de la propuesta, que necesariamente contendrá elementos de temas didácticos y de la práctica profesional que se desarrolla en el aula. Por otro lado, una propuesta de Apoyo a la docencia, puede tener múltiples participantes de diversas disciplinas y niveles, que brindarán apoyo a la actuación del profesor y tratarán de mejorar el entorno educativo (Barraza, 2010).

El mismo autor describe las fases y momentos que la constituyen:

- Fase de planeación: Consiste en la elección del tema de interés, la construcción de la problemática y el diseño de la solución.
- Fase de implementación: Es el momento de la aplicación de las actividades definidas en la propuesta de intervención y su reformulación en caso de requerir una adaptación.

- Fase de evaluación: Se realiza el monitoreo de la aplicación de las diferentes actividades indicadas en el proyecto y su evaluación general. Es importante el seguimiento durante todo el proceso para corregir eventualidades.
- Fase de socialización-difusión: En ella se socializan los resultados y se invita a los involucrados en la problemática a adoptar y recrear la solución diseñada.

### **3.5 Educación Mediada por Tecnología**

El concepto de Tecnología Educativa (TE) ha evolucionado a la par del proceso de maduración de la tecnología y al desarrollo de los distintos enfoques teóricos de la enseñanza y el aprendizaje. Sin embargo, es gracias al auge de las tecnologías de la información y la comunicación (TIC) que se asume como una disciplina pedagógica de reflexión y teorización de la actividad educativa planificada de acuerdo con sus contextos (Torres y Cobo, 2017). Se entiende la tecnología educativa como una forma de planificar y poner en práctica la educación, configurando los procesos de enseñanza y aprendizaje, sus recursos, espacios y tiempos, en función de intencionalidades bien definidas (Sancho et al., 2015).

De acuerdo con Serrano et al. (como se citó en Torres y Cobo, 2017, p. 33), se afirma que:

La tecnología educativa constituye una disciplina encargada del estudio de los medios, materiales, portales web y plataformas tecnológicas al servicio de los procesos de aprendizaje; en cuyo campo se encuentran los recursos aplicados con fines formativos e instruccionales, diseñados originalmente como respuesta a las necesidades e inquietudes de los usuarios (...) radica en un enfoque sociosistémico, donde ésta siempre analiza procesos mediados con y desde una perspectiva holística e integradora.

Como resultado del desarrollo de las TIC, muchos de los paradigmas sobre la educación han ido cambiando. La incorporación de nuevas formas de aprendizaje, incluyendo una gran variedad de contenidos y formas de interacción, ha sido posible gracias al uso de sistemas de aprendizaje automatizados, Internet y sistemas Web; así como de nuevos dispositivos electrónicos de comunicación, como los dispositivos móviles y los teléfonos inteligentes. El acceso a los contenidos de aprendizaje basados Internet y redes de computadoras, ha propiciado nuevas modalidades de enseñanza y aprendizaje, como el *e-learning* definido como una forma de aprendizaje basado en tecnología electrónica (Ji et al., 2016), o el *b-learning* (*blended learning*), que combina el trabajo presencial en aula con el trabajo en línea mediante Internet y recursos digitales.

Cabero (2016), en la revisión de diversas investigaciones sobre TE, tendencias y proyecciones futuras, destaca algunos aspectos a considerar al momento de emprender nuevas investigaciones y aplicaciones de la TE: superar el modelo comparativo de medios, en el que cuando aparece una nueva tecnología se plantea de inmediato si esta es mejor o más eficaz que las existentes para ser utilizada en la educación; controlar el efecto de la novedad; superar el solucionismo tecnológico, ya que la incorporación de una nueva tecnología no siempre garantiza mejoras en el aprendizaje; y recordar que el docente es quien finalmente y a su modo controla la aplicación de los medios tecnológicos en el proceso educativo.

### **3.5.1 Ambientes Virtuales de Aprendizaje**

El Sistema de Gestión de Aprendizaje (LMS), también denominado Sistema de Gestión de Cursos (CMS) o Ambiente Virtual de Aprendizaje (VLE), ha sido utilizado desde hace décadas y ha evolucionado para convertirse en parte fundamental de la infraestructura tecnológica para la instrucción en instituciones educativas de todos los niveles (Rhode et al., 2017). Trabajos recientes con estudiantes de educación superior para conocer sus preferencias tecnológicas, han señalado que casi todos

los estudiantes usan un LMS y lo colocan como una de las tecnologías educativas más importantes para su éxito académico (Brooks, 2016).

El aprendizaje llevado a cabo en entornos en línea requiere de la gestión de los alumnos y sus contenidos. Un LMS es una herramienta para gestionar los procesos, el historial y los contenidos en un entorno de aprendizaje electrónico y para facilitar un aprendizaje continuo. Las principales funciones que las herramientas de un LMS permiten concretar son: la administración del curso, la comunicación entre los participantes, la gestión de contenidos, la gestión del trabajo en grupos, y la evaluación (Fernández-Pampillón, 2009).

Estos sistemas se han utilizado en una amplia variedad de entornos, desde escuelas de educación básica hasta universidades, empresas e instituciones públicas. Algunos de los sistemas LMS típicos son Moodle, Blackboard, Dokeos, Shakai, e-College y Claroline, unos pueden ser utilizados de manera libre mientras que otros requieren la adquisición de licencias.

Los MOOC (*Massive, Open, Online, Courses*) son una modalidad de formación abierta y gratuita basada en sistemas en línea con materiales de cursos creados por universidades y entregados generalmente en formato de video y texto. Tienen un enfoque de autoaprendizaje que le permite al alumno controlar dónde y cuándo aprender, pueden hacer pausas y reflexionar sobre los contenidos, aunque, sin el diseño e inclusión de mecanismos de interacción, puede ser una experiencia de aprendizaje en solitario (Sharples et al., 2014). Por ello la necesidad de involucrar a los estudiantes mediante herramienta y estrategias que faciliten el aprendizaje social y colaborativo, al compartir sus ideas y discutir sus diferentes perspectivas a medida que aprenden en línea, y realizar evaluaciones integrales; a este tipo de sistemas se les ha denominado cMOOC, cuya “c” viene de “conectivista” (Bates, 2015). Algunas de las plataformas más populares son Coursera, Udacity, edX, FutureLearn, Miríadax, entre otras.

Otra variante son los denominados cursos SPOC (*Small Private On-Line Course*), que corresponde a una adecuación de los MOOC para atender necesidades específicas de instituciones educativas y grupos concretos de usuarios, de tal suerte que los estudiantes pueden aprovechar todas las potencialidades de los MOOC, pero desarrollando perfiles educativos específicos acorde a sus necesidades particulares (Fox et al., 2014).

El reto que se vislumbra en los trabajos analizados, es el diseño instruccional apropiado acorde a los objetivos de aprendizaje de los cursos, para aprovechar de manera eficiente e integrada las tecnologías y recursos disponibles en las plataformas de gestión de ambientes virtuales de aprendizaje. Otro aspecto a considerar es la selección de la tecnología apropiada y el desarrollo de habilidades para su empleo por parte de los instructores (Bates, 2015).

### **3.6 Diseño Instruccional**

La puesta en marcha de la acción educativa debe ser precedida por el diseño cuidadoso de la planeación didáctica en la que se integran los elementos teóricos y prácticos relacionados con la enseñanza y el aprendizaje, en donde además se deben especificar los recursos materiales y tecnológicos, acordes a las necesidades de los estudiantes. El objetivo del diseño instruccional es definir los elementos necesarios para el óptimo desarrollo del proceso educativo, en el que deben estar alineados los recursos y las estrategias utilizadas para alcanzar las metas de aprendizaje, de acuerdo con las necesidades de los aprendices y los elementos del contexto educativo.

#### ***3.6.1 Perspectiva Teórica del Diseño Instruccional***

Desde un enfoque cognoscitivo, los procesos mentales para representar y utilizar el conocimiento son primordiales, especialmente es de interés el describir y explicar la naturaleza de las representaciones mentales, así como el determinar el papel que juegan en la producción de las acciones y las conductas humanas (Pozo, 2008) .

La enseñanza está orientada al logro de habilidades y de estrategias para alcanzar un aprendizaje eficaz, además de aplicar los conocimientos adquiridos ante situaciones de cualquier tipo. Por ello, se hace énfasis en el desarrollo de actividades que fomenten el aprendizaje significativo y a su vez, que hagan uso de estrategias para el autoaprendizaje y la resolución de problemas, además de habilidades metacognitivas que les permitan a los alumnos desenvolverse en contextos y situaciones nuevas. Por tanto, en el paradigma cognitivo, aprender es producto del uso efectivo de las estrategias cognitivas, metacognitivas y autorregulatorias (Hernández Rojas, 1997).

Las estrategias de aprendizaje son planes o cursos de acción que sirven para optimizar los procesos cognoscitivos y para la realización exitosa de una tarea. Requieren de la participación de otras estrategias denominadas metacognoscitivas, que sirven para entender que habilidades, conocimientos y recursos que se necesitan al realizar una tarea, pero también, son necesarias para saber cómo y cuándo se deben de utilizar dichos recursos. Las estrategias autorregulatorias están estrechamente relacionadas a las anteriores, y son entendidas como un proceso mediante el cual el alumno establece sus propios objetivos de aprendizaje, y realizan tareas de supervisión, regulación y control de sus pensamientos, sus emociones y su motivación para dirigir sus acciones hacia las metas (Escamilla y Heredia, 2019; Schunk, 2012).

De acuerdo con la teoría cognoscitiva social:

Las metas y las expectativas son mecanismos de aprendizaje importantes. La motivación es la conducta dirigida a metas que es instigada y mantenida por las expectativas de las personas en relación con los resultados que anticipan de sus acciones y de su autoeficacia para realizar esas acciones (Schunk, 2012, p. 372).

Los aprendices motivados realizan tareas autorregulatorias como elementos de ayuda para lograr sus metas, por ejemplo, se monitorean, organizan sus tareas y

sus tiempos, y ajustan sus estrategias cuando consideran que es necesario; lo que contribuye a un mejor aprendizaje. La motivación está relacionada recíprocamente con la autorregulación (Schunk, 2012).

Como se ha mencionado, para la realización efectiva de los procesos cognitivos se depende en buena medida de que existan las condiciones adecuadas para su buen funcionamiento. Algunos de los procesos auxiliares que contribuyen a la eficacia del aprendizaje son: la motivación, la atención, la recuperación de aprendizajes anteriores y la metacognición, como un mecanismo para controlar o autorregular el propio aprendizaje.

### ***3.6.2 Modelo de Diseño Instruccional Mediado con Tecnología***

La propuesta de un curso deberá establecer las mejores condiciones para llevar a cabo la enseñanza de manera que se favorezcan los procesos de aprendizaje, analizando y adecuando los recursos instruccionales y tecnológicos de acuerdo con las características y necesidades educativas de los estudiantes. Se trata, al final de cuentas, de maximizar los procesos cognitivos y sus procesos auxiliares. Un buen diseño instruccional (DI) tratará de alinear los componentes del aprendizaje para que todos apunten a la misma dirección.

El propósito del diseño instruccional, en un contexto educativo mediado por tecnología, es obtener el máximo aprovechamiento pedagógico de la tecnología y de las herramientas multimedia. Es una propuesta que busca confirmar con anticipación si un proceso de formación será eficaz didáctica y económicamente. En contextos de educación virtual, el DI contempla el uso adecuado de estrategias y recursos que aseguran una interactividad y calidad en la instrucción, para el logro de las metas de aprendizaje.

Un proceso de DI describe los detalles del desarrollo, implementación, evaluación y mantenimiento de escenarios que facilitan el aprendizaje en diversos niveles de complejidad o de magnitud de unidades de contenido. También, se deben

contemplar las necesidades de aprendizaje y el contexto en que se manifestarán, por lo que requiere de objetivos de formación concretos para seleccionar los recursos más adecuados, y desarrollar contenidos, actividades y evaluaciones apropiadas.

El diseño instruccional seleccionado es el modelo ADDIE, cuyas características se describen con detalle en Williams et al. (2019); incluye una adaptación del diseño a cursos basados en la Web. El acrónimo ADDIE indica los pasos clave del proceso: *Analysis* (análisis), *Design* (diseño), *Development* (desarrollo), *Implementation* (implementación) y *Evaluation* (evaluación). Estos pasos pueden seguirse secuencialmente, o también de manera ascendente y simultánea a la vez.

### **3.7 Minería de Datos Educativos**

El desarrollo de herramientas y algoritmos computacionales para extraer información potencialmente útil y novedosa a partir de bases de datos generadas en contextos educativos ha dado lugar a una nueva disciplina emergente de las ciencias computacionales denominada Minería de Datos Educativos (MDE, o EDM por sus siglas en inglés *Educational Data Mining*). La cual es un área de investigación interdisciplinaria que maneja el desarrollo de métodos para explorar los datos que surgen en un campo escolar. Los enfoques computacionales utilizados por la MDE permiten examinar y estudiar datos escolares para responder a preguntas educativas.

La MDE está basada en conceptos y técnicas de la Minería de Datos (MD), tiene el propósito de explorar datos provenientes de entornos educativos para encontrar patrones descriptivos y predicciones que caracterizan los comportamientos y logros de los alumnos, el contenido de conocimiento de dominio, evaluaciones, funcionalidades educativas y aplicaciones; empleando para ello diversos métodos, técnicas y algoritmos de descubrimiento de información (Peña-Ayala, 2014).

La Analítica del Aprendizaje (AA, sus siglas en inglés son LA, por *Learning Analytics*), es una disciplina muy relacionada con la MDE, que realiza tareas de medición, recopilación, análisis y presentación de datos de los alumnos y sus contextos, a efectos de entender y optimizar el aprendizaje y los entornos en los que ocurren los sucesos de aprendizaje.

La MDE proporciona un conocimiento intrínseco del proceso de enseñanza y aprendizaje, que, mediante la generación de modelos, pueden responder a preguntas relacionadas con el rendimiento escolar y sus problemáticas asociadas y facilitar una planificación educativa efectiva (Anoopkumar y Rahman, 2016).

El propósito de estas técnicas es generar información para retroalimentar el proceso de enseñanza-aprendizaje de alumnos, profesores y administrativos. Sin embargo, no siempre es sencillo entender los resultados de los análisis obtenidos mediante minería de datos, se requiere de grandes esfuerzos para usuarios no especializados que no están familiarizados con los algoritmos y las técnicas empleadas (Ji et al., 2016).

### **3.7.1 Tareas de la Minería de Datos Educativos**

La minería de datos educativos ofrece formas alternativas de analizar los datos escolares y descubrir información relevante en las poblaciones estudiadas y los entornos de aprendizaje. A continuación, se describen algunas de sus principales aproximaciones o métodos de análisis de acuerdo con las técnicas que emplean (Peña-Ayala, 2014):

- Predicción: desarrollo de modelos para inferir una variable a partir de la combinación de otros datos disponibles o variables predictoras. Algunos de los métodos usados para predecir son: *clasificación* (cuando la variable a predecir es un valor categórico), *regresión* (si la variable tiene valor continuo), o *estimación de densidad* (cuando la variable a predecir es una función de

densidad de probabilidad). Se pueden aplicar estos métodos para predecir el éxito académico de los estudiantes.

- Agrupamiento: para encontrar conjuntos de datos que se agrupen naturalmente, y separarlos del conjunto completo en una serie de categorías. Se pueden crear grupos de estudiantes basados en sus patrones de aprendizaje o estrategias cognitivas, por ejemplo.
- Minería de relaciones: se usan para descubrir relaciones entre variables y codificarlas como reglas para aplicarlas posteriormente. Algunos de sus métodos son: minería de reglas de asociación (cualquier relación entre variables), minería de patrones secuenciales (asociaciones temporales entre variables), minería de correlaciones (correlaciones lineales), minería de datos causales (relaciones causales entre variables). Usadas para identificar relaciones, de las actividades de los estudiantes con sus resultados finales, y para modelar secuencias de actividades de aprendizaje.
- Descubrimiento mediante modelos: usado para validar el modelo de un fenómeno (mediante predicción, agrupamiento o ingeniería del conocimiento). Es usado para identificar la relación entre características y comportamiento de los estudiantes.
- Destilado de datos: para permitir a un humano identificar o clasificar rápidamente las propiedades de un conjunto de datos. Utiliza resúmenes, visualización e interfaces interactivas para destacar la información relevante y apoyar la toma de decisiones.
- Detección de valores atípicos: para descubrir datos significativamente diferentes al resto del conjunto. Se puede utilizar para detectar desviaciones en las acciones o comportamientos del alumno o educador, procesos de aprendizaje irregulares y para detectar estudiantes con dificultades de aprendizaje.

Con el tiempo se han sumado más metodologías al repertorio de la MDE, por ejemplo, el Análisis de Redes Sociales (SNA), la Minería de Textos, la Minería de Procesos y el Trazado de Conocimientos, entre otras. La elección de una técnica determinada depende del entorno educativo, los objetivos de investigación y la disponibilidad de los datos. La MDE ofrece una forma novedosa de hacer investigación educativa, respecto a los paradigmas más tradicionales. Incluye métodos basados en estadística, probabilidad, aprendizaje automático, redes neuronales, entre otros, que permiten distintos tipos de análisis; también dispone de mecanismos para medir la precisión de los resultados y los modelos generados por sus diversas técnicas.

### **3.8 Instrumentos Teóricos de Análisis de Datos**

Como principal instrumento para el análisis de los datos se consideraron las diversas técnicas de la minería de datos educativos, que incluyen algoritmos para analizar datos e identificar patrones, con lo cual se pueden obtener modelos de clasificación y predicción, entre otros. Así mismo, incluyen diversos métodos para validar los resultados y garantizar niveles adecuados de precisión en las predicciones.

#### **3.8.1 Técnicas de Minería de Datos Educativos**

Dada la naturaleza de los estudios sobre rendimiento académico o desempeño académico utilizando MDE, se han empleado principalmente métodos de clasificación, regresión, agrupamiento y minería de relaciones, siendo la tarea de clasificación la más popular (Shahiri et al., 2015).

Los algoritmos ubicados en estas categorías de tareas son de naturaleza inductiva, es decir, buscan derivar o descubrir inductivamente características o patrones a partir de los datos (datos de entrenamiento). El conocimiento obtenido se sintetiza en modelos predictivos los cuales pueden emplearse con nueva información (datos de prueba). Se definen dos grandes grupos de tipos de algoritmos (Minguillón et al., 2017):

- Métodos supervisados: que realizan su proceso de aprendizaje con base en un conjunto de datos en donde los valores de las clases son conocidos.
- Métodos no supervisados: cuando no se conocen los valores de las clases o no están definidos de antemano.

Los principales métodos de minería de datos predictivos de tipo supervisado, de acuerdo con los mismos autores, son: k-NN, redes neuronales, árboles de decisión, SVM y métodos probabilísticos.

Algunos de los métodos/técnicas más utilizados por su efectividad para la predicción se describen a continuación:

- **Árbol de decisión (AD):** es un método de aprendizaje supervisado que construye un árbol a partir de un conjunto de datos históricos (datos de entrenamiento), encontrando patrones predictivos inductivamente. Cada rama del árbol es una posible explicación al problema de clasificación o predicción planteado (reglas de decisión o clases). Las reglas son un conjunto de condicionales Si-Entonces y se interpretan leyendo la secuencia de nodos y sus condiciones partiendo de la raíz, hasta llegar al nodo hoja (Peña-Ayala, 2014).
- **Redes neuronales (RN):** tienen la capacidad de detectar las interacciones entre variables predictoras, incluso relaciones no lineales entre variables dependientes e independientes. Son modelos que imitan a sistemas biológicos, utilizan un grafo dirigido compuesto por neuronas que se interconectan a través de arcos dirigidos, que establecen la relación entre los nodos. El arco tiene un peso numérico asociado para indicar dirección y fuerza de la relación. Existen niveles o capas que son conjuntos de neuronas clasificadas como: entrada, oculta o salida. Se definen dos funciones: una función de activación y una función de salida. El modelo de RN más conocido, llamado Perceptrón Multicapa, utiliza una capa de entrada, una oculta y una de salida. Los nodos de entradas se asocian a cada una de las variables

predictoras, los nodos intermedios son interconexiones y los nodos de salida se asocian a los valores de la variable dependiente. La RN hace un proceso de inferencia para predecir la clasificación del resultado final del rendimiento académico, calculando los costos al realizar la propagación de la red (Menacho, 2017).

- Redes bayesianas (RB): son redes probabilísticas que emplean un grafo para representar un modelo causal, en donde los nodos son las variables dependientes e independientes. Los arcos del grafo indican las relaciones de dependencia entre variables. Se utilizan valores de probabilidad condicional para reflejar posibles relaciones causa – efecto entre los nodos. Para problemas de predicción y clasificación es muy utilizada la RB conocida como Naive Bayes; en la cual, el nodo raíz representa a la variable dependiente (variable clase) y el conjunto de nodos restantes son las variables independientes. Este tipo de red emplea aprendizaje paramétrico, ya que requiere estimar las probabilidades desde el conjunto de entrenamiento. El proceso de predicción o clasificación de la RB se basa en estimar la probabilidad a posteriori, para determinar la clase más probable a la que pueda pertenecer la variable dependiente (Miranda y Guzmán, 2017).
- Algoritmo k-NN: el método de “k vecinos más cercanos” (en inglés, *k nearest neighbours*) clasifica una nueva instancia de datos y realiza su aprendizaje de forma supervisada de forma simultánea. Por cada dato de entrada, se calcula la distancia con todas las instancias de entrenamiento y se eligen las *k* más próximas, según alguna métrica de similitud. El valor de la clase de la nueva instancia queda determinado por la vecindad con las *k* instancias más cercanas, el valor de *k* es inicialmente pequeño (Minguillón et al., 2017). Una desventaja de este método es que el coste computacional de calcular una nueva clasificación puede crecer rápidamente si se tienen muchos datos de entrenamiento o si hay demasiados atributos.

### **3.8.2 Validación de Modelos Predictivos**

Con el fin de implementar el modelo predictivo y que sea de utilidad práctica, es necesario validarlo de manera que sus resultados o pronósticos sean confiables, es decir, comprobar si el modelo creado funciona igual con otros individuos distintos a los que se emplearon para crearlo. Para ello, se realiza la validación utilizando las opciones integradas en el software de minería de datos, en donde se puede reservar un porcentaje de casos de la muestra para aplicar pruebas de validación cruzada y también, al ejecutar los algoritmos, se obtienen valores estadísticos que miden la precisión de las variables y los resultados.

Para evaluar la precisión de los modelos obtenidos mediante las técnicas de clasificación, en primer lugar, se debe aplicar el algoritmo y obtener la clasificación o predicción de la clase, luego comparar con el valor real de las instancias y determinar la precisión y errores de la clasificación.

La precisión es el número de predicciones correctas, tanto para una clase positiva como para la clase negativa, sobre el número total de predicciones:

$$\text{Precisión} = [(VP + VN) / N] \quad (1)$$

En donde VP es verdadero positivo, VN es verdadero negativo y N la cantidad de instancias.

El error indica la tasa de mala clasificación y está determinado por:

$$\text{Error} = [(FP + FN) / N] \quad (2)$$

En donde FP es falso positivo, FN es falso negativo y N la cantidad de instancias.

Una manera alternativa de evaluar el rendimiento de un modelo generado con algoritmos de aprendizaje máquina es mediante las curvas ROC (*Receiver Operating Characteristics*) (Fawcett, 2003), la cual es una técnica para visualizar su desempeño y ha sido adoptada por la comunidad de investigadores en minería de

datos. Se calcula con base en los falsos positivos (FP) y los verdaderos positivos (VP). A partir de la curva ROC se calcula el AUC (*Area Under the Curve*), que es un indicador de la calidad del modelo. Si el valor de AUC se acerca a la unidad, significa que el clasificador se comporta de manera óptima, aproximándose al clasificador perfecto.

De acuerdo con (Minguillón et al., 2017), se pueden interpretar las curvas ROC utilizando los intervalos de valores AUC de la siguiente guía:

[0.5 – 0.6): Test malo

[0.6 – 0.75): Test regular

[0.75 – 0.9): Test bueno

[0.9 – 0.97): Test muy bueno

[0.97 – 1]: Test excelente

El software WEKA cuenta con las siguientes opciones para el proceso de evaluación y validación de modelos (Bouckaert et al., 2018):

- Utilizar el conjunto de entrenamiento: se evalúa el clasificador utilizando el mismo conjunto de datos utilizado para construir el modelo, por lo que el resultado suele ser demasiado optimista.
- Suministrar un conjunto de prueba: se consigue un conjunto de datos independiente para realizar las pruebas. Para cada instancia del nuevo conjunto se realiza la predicción y se estima la precisión del modelo.
- Validación cruzada (*k-fold cross validation*): se divide el conjunto de datos en K partes (*folds*) y se realizan K iteraciones en las que se va reservando una parte como datos de prueba y el resto como datos de entrenamiento. Al final se promedian los resultados de las precisiones y errores calculados en todas las iteraciones.

- Indicar un porcentaje de división: se dividen los datos en un grupo de entrenamiento y un grupo de prueba, de acuerdo con un porcentaje indicado.

La validación cruzada es la forma más consistente de evaluar clasificadores, ya que los conjuntos se determinan de forma aleatoria y el error de la evaluación es muy bajo. Comparado con los demás métodos, es la forma más estricta de verificar la precisión de los modelos, esto asegura su capacidad de generalización.

Por lo anterior, para la validación de los modelos predictivos del proyecto, se ha considerado el empleo de validación cruzada. Y para identificar la calidad de los mismos, se utiliza el indicador AUC-ROC.

## **4. METODOLOGÍA**

En este capítulo se presentan los elementos metodológicos empleados en el estudio. Se plantea la hipótesis, los supuestos de la investigación y los objetivos perseguidos. Además, se especifica el enfoque adoptado para el desarrollo de la investigación, específicamente se hace una descripción general de la metodología mixta empleada y de las etapas que fueron definidas para el desarrollo de las actividades. Particularmente, se describen los procesos de investigación-acción y de la minería de datos educativos. También, se delimita el alcance, los recursos, la población y las estrategias de recolección de datos de la investigación. Finalmente, se mencionan los distintos procedimientos de recolección y análisis de datos empleados, cuya variedad de información generada permite tener una visión más amplia del fenómeno estudiado

### **4.1 Determinación de Hipótesis y Premisa**

A partir de las preguntas de investigación y considerando la etapa cuantitativa de la tesis se formuló la siguiente hipótesis:

El rendimiento escolar de alumnos de primer ingreso a carreras del área de computación en la Facultad de Matemáticas de la Universidad Autónoma de Yucatán, está relacionado con antecedentes extraescolares y académicos de los estudiantes, que al ser analizados pueden servir para generar modelos predictivos eficientes.

En un plano general, el supuesto de la investigación es que la información, los modelos y sistemas generados durante el desarrollo del proyecto de investigación, sirven como instrumentos estratégicos para la planificación educativa y toma de decisiones oportuna de tutores, docentes, directivos o alumnos y esto impacta de manera positiva en el rendimiento académico de los estudiantes.

## **4.2 Objetivos del estudio**

### **4.2.1 Objetivo General**

Desarrollar un sistema predictivo de rendimiento escolar de alumnos de primer ingreso a carreras del área de computación en la Facultad de Matemáticas de la UADY, utilizando modelos generados mediante minería de datos educativa, para la detección y atención oportuna de problemas asociados al bajo rendimiento académico.

### **4.2.2 Objetivos Específicos**

OE1. Diseñar modelos predictivos eficientes de rendimiento escolar con base en los antecedentes extraescolares y académicos de los estudiantes, datos recuperados de los sistemas de información institucional.

OE2. Implementar modelos predictivos mediante un sistema de software que facilite el análisis de nuevas instancias de alumnos y genere informes de predicción de riesgo académico.

OE3. Identificar elementos del contexto académico que puedan afectar la transición de los estudiantes a la universidad y que condicionen su desempeño académico.

OE4. Definir estrategias de intervención educativa para los alumnos con probable bajo rendimiento académico, con énfasis en la utilización de tecnología educativa.

OE5. Evaluar la efectividad de las estrategias de intervención implementadas en la búsqueda de la mejora continua de los procesos educativos.

## **4.3 Tipo de Investigación**

Se trata de un trabajo con enfoque de investigación mixto, tanto cuantitativo como cualitativo, llevado a cabo en etapas o fases (Hernández et al., 2014).

La primera etapa se desarrollada con el enfoque cuantitativo, empleando un diseño no-experimental (*ex post facto*), en el cual la selección de los sujetos es posterior a

haber ocurrido el fenómeno y se busca reconstruir las posibles causas que lo han ocasionado sin manipular las variables (Ballester et al., 2017). Específicamente, se utilizan los procedimientos de la Minería de Datos Educativos para generar modelos predictivos de rendimiento escolar que posteriormente se implementan en un sistema informático.

En la segunda etapa, con una orientación de investigación cualitativa, se determinan y aplican estrategias de intervención académica para los alumnos con problemas de rendimiento académico deficiente, para ello se parte de una perspectiva de investigación-acción (Álvarez y Álvarez, 2014).

Un aspecto importante de la investigación, que da sustento teórico al desarrollo de la misma, es la revisión de la literatura, por lo que, para la elaboración del estado del arte y marco teórico, se trabajó con un enfoque cualitativo e interpretativo, de tipo documental (Gómez et al., 2015).

El desarrollo del sistema predictivo se realizó siguiendo las pautas de la metodología de desarrollo ágil (Orjuela y Rojas, 2008), la cual permite definir de forma rápida, los requisitos y funcionalidades del sistema, diseñar la arquitectura que gestionará los componentes de acceso a WEKA y Moodle, desarrollar los componentes de software, realizar pruebas de funcionamiento, evaluar su desempeño y, finalmente, utilizar el sistema en la predicción de nuevos casos de alumnos de recién ingreso.

#### ***4.3.1 El Enfoque de la Metodología Mixta***

El diseño de investigación con métodos mixtos es un procedimiento para recopilar y analizar datos combinando los métodos cuantitativos y cualitativos en un solo estudio o una serie de estudios para comprender un problema de investigación. Se asume que el uso armonizado de ambos métodos resultará en un mejor entendimiento del problema y responderá mejor a las preguntas de investigación (Creswell, 2012).

Sin embargo, es necesario entender que el método mixto consiste en fusionar, integrar, vincular o incrustar las secuencias o flujos de trabajo de la investigación, por lo que los datos cuantitativos y cualitativos son analizados en conjunto para obtener los resultados finales, que deberán de responder a las preguntas o hipótesis de la investigación.

Algunas de las funciones y justificantes del empleo de los métodos mixtos, de acuerdo a Hernández et al. (2014), son: incrementar la validez mediante la triangulación de resultados, compensar o complementar las debilidades de los métodos, examinar los procesos más holísticamente, responder a múltiples preguntas de investigación y más ampliamente, usar los resultados de un método para generar hipótesis que pueden ser confirmadas por otro método, generar nuevos métodos de recolección y análisis de datos, entre otras.

Existen algunas variantes de diseños de métodos mixtos (Creswell, 2014; Hernández et al., 2014). Considerando las preguntas, objetivos, complejidad y alcance de este trabajo de investigación, se determinó adoptar el método mixto de múltiples fases, el cual es común en los campos de evaluación y programas de intervención. En este diseño, se utilizan estrategias simultáneas o secuenciales a lo largo del tiempo para comprender mejor los procesos estudiados.

En la Figura 1, se ilustra un diseño de investigación utilizando el método mixto con enfoque de múltiples fases.

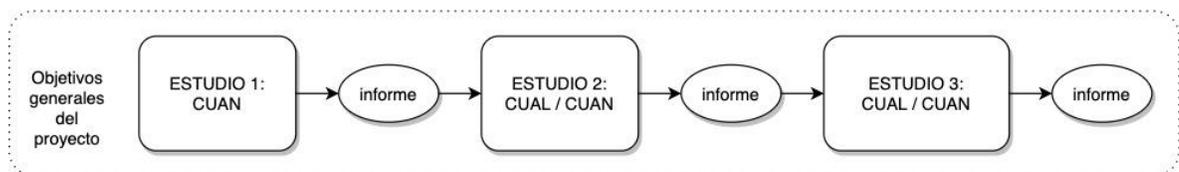


Figura 1 Caso de diseño de múltiples fases de la metodología mixta.

Fuente: Adaptación a partir de (Creswell, 2012).

Para la comparación y análisis de los resultados finales, lo más común es describir los resultados cuantitativos y cualitativos uno al lado del otro en la sección de discusiones del informe. Por ejemplo, presentar los resultados estadísticos cuantitativos y luego proporcionar citas cualitativas para confirmar o no los resultados estadísticos.

#### **4.4 Proceso de Investigación-Acción**

De acuerdo con Ruiz-Bernardo et al. (2018) se entiende la investigación–acción (I-A) como una forma de indagación autorreflexiva y correflexiva llevada a cabo por un grupo de personas con intereses y preocupaciones en común, que mediante acciones planificadas tratan de resolver o mejorar situaciones problemáticas en la cotidianidad de sus vidas; en este caso, en el ámbito de la educación superior.

La información que surge durante la implementación de las etapas de la metodología sirve de guía para la toma de decisiones en la mejora continua de programas, procesos y cambios estructurales en los entornos donde se aplica.

La investigación-acción se desarrolla a través de una espiral de ciclos, durante los que se suceden fases de planificación, acción, observación y reflexión. Se parte de la identificación de un problema inicial, del que se recogen evidencias y analizan los datos obtenidos, a partir del cual se desarrolla un plan de acción flexible, al que sigue la acción y reflexión, derivando en la planificación de un nuevo ciclo (Ruiz-Bernardo et al., 2018, p. 37).

En la Figura 2 se representan las etapas principales del proceso. A partir de la reflexión y retroalimentación es posible volver a iniciar nuevos ciclos del proceso pues lo que se busca es la mejora continua mediante la detección de problemáticas y la realización de actividades planificadas que intenten resolverlas.

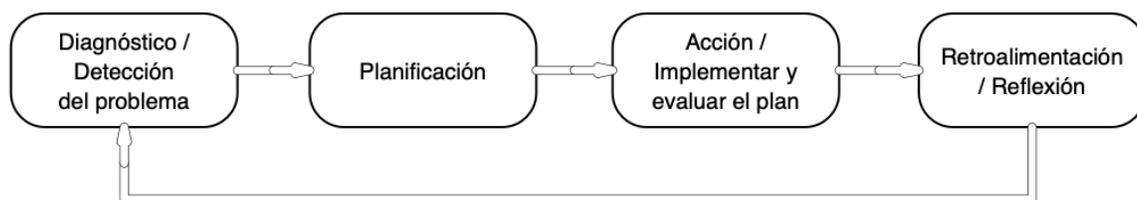


Figura 2 Fases del proceso de Investigación-Acción.

Fuente: Adaptación a partir de Ruiz et al. (2018).

#### 4.5 Proceso de la Minería de Datos Educativos

Para el empleo de la MDE se utilizarán las técnicas de clasificación para obtener modelos predictivos, y se complementará con la implementación de un sistema informático, mediante el cual se espera analizar la situación académica de los alumnos que recién ingresan a la Institución y detectar probables casos de bajo rendimiento, con ello se dará oportunidad de aplicar medidas de intervención educativa de manera oportuna.

Específicamente, el desarrollo de la investigación se sustenta en el proceso de la minería de datos educativos para generar modelos predictivos de rendimiento académico y riesgo escolar, que posteriormente serán implementados en un sistema informático. Se han formulado varios esquemas para aplicar la MDE basados en la metodología conocida como KDD (*Knowledge Discovery in Databases*) o Proceso de Extracción de Conocimiento en Bases de Datos (Márquez-Vera, Romero y Ventura, 2012; Peña-Ayala, 2014).

Una propuesta del proceso de la minería de datos educativos se muestra en la Figura 3. Los pasos indicados permiten descubrir conocimiento que da respuesta a cuestionamientos concretos de los entornos educativos, que proporcionan información para la toma de decisiones, además, en cada ciclo se van refinando los criterios para obtener mejores resultados.

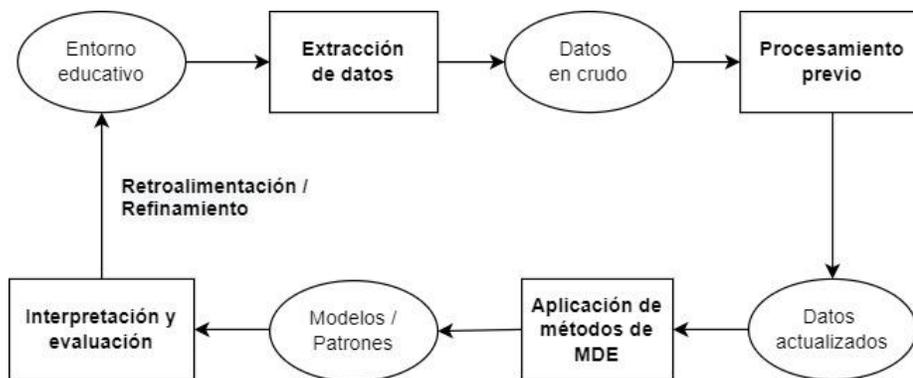


Figura 3 Proceso de la minería de datos educativa.

Fuente: Adaptación a partir de Alyahyan y Düşteğör (2020) y Anoopkumar y Rahman (2016).

Hay una serie de tareas esenciales en el proceso de extracción de conocimiento y la aplicación de la MDE, como los describen varios autores (Anoopkumar y Rahman, 2016; Merchan y Duarte, 2016; Peña-Ayala, 2014). Una descripción breve de estas tareas se presenta a continuación:

- Extracción de datos: consiste en la selección de los datos de estudio del entorno educativo y su recuperación. Las fuentes de datos pueden ser heterogéneas: clases tradicionales, cursos *e-learning*, sistemas basados en web, redes sociales, entre otras. Se da preferencia a datos en formato digital provenientes de sistemas de información o entornos virtuales de aprendizaje, por conveniencia. Los datos pueden quedar agrupados en bases de datos en un formato crudo, es decir, sin ningún tratamiento aún. La información recuperada puede corresponder a la interacción y uso de los estudiantes en los sistemas, información de cursos, datos académicos, o datos provenientes de encuestas socioeconómicas, antecedentes escolares, entre otros.
- Procesamiento previo: una vez almacenados los datos en crudo, se limpian o eliminan los elementos con valores redundantes, incorrectos o faltantes, se fusionan datos provenientes de múltiples fuentes y se convierten valores de datos a formas más apropiadas para su manejo o análisis, es decir, se validan los datos para un manejo estandarizado y homogéneo. Esta etapa, por su

complejidad, suele necesitar la aplicación de técnicas específicas para obtener datos correctamente estructurados. Es recomendable excluir variables o atributos irrelevantes, esto es, atributos que no aportan información útil para resolver el problema. Es posible que se requiera convertir los archivos a formatos específicos del sistema de minería de datos a utilizar. El resultado de esta etapa es un conjunto de datos actualizado y listo para su análisis.

- Aplicación de métodos de MDE: consiste en la selección y aplicación de las técnicas más adecuadas acorde a las necesidades o hipótesis del estudio. Por los objetivos perseguidos en este trabajo, se emplearán métodos de clasificación/predicción para construir modelos que identifiquen casos de riesgo académico de los alumnos. Una vez cargados los datos, es posible seleccionar atributos para ser considerados en el análisis, según su relevancia. Para ello, pueden emplearse filtros que facilitan la elección de atributos concretos, además de la definición de los parámetros de funcionamiento de los algoritmos, con lo que se puede ir refinando la precisión de los modelos generados.
- Interpretación de resultados: la salida resultante de los algoritmos se muestra para su interpretación, los modelos muestran los hallazgos de patrones, relaciones o tendencias en los datos. Se analizan los modelos con los mejores resultados para determinar el riesgo escolar de los estudiantes. También se pueden analizar los valores generados en la medición de la precisión de los resultados y utilizar métodos de validación. Dependiendo de la efectividad o validez de los reglas y modelos pueden requerirse más iteraciones para refinar el funcionamiento de los algoritmos (incluir otros atributos, remover atributos “ruidosos” o que no tienen peso en el modelo, aplicar filtros, etc.).
- Retroalimentación/Refinamiento: el conocimiento obtenido es informado a los actores del entorno educativo, a su vez, se pueden incluir recomendaciones;

con lo cual es posible planificar acciones para modificar o mejorar procesos o condiciones del contexto escolar. También, en este proceso iterativo, se pueden ajustar diversos elementos para obtener modelos más confiables, al incorporar nuevas variables o datos del entorno educativo, e incluir información de nuevos grupos de estudiantes, con lo cual se pueden mejorar los modelos predictivos.

#### **4.6 Alcance de la Investigación**

La investigación tiene un enfoque de trabajo mixto de múltiples fases, tanto cuantitativas como cualitativas, por lo que para cada una de las etapas se establecen alcances específicos que, en conjunto, se complementan para lograr los propósitos generales de la investigación. En la primera etapa, se considera un alcance correlacional ya que se analiza cómo las variables o factores académicos o no académicos influyen en el rendimiento escolar. En este caso, la información requerida son datos académicos de antecedentes escolares y extraescolares de los estudiantes de primer semestre que hayan ingresado entre 2016 y 2018. Posteriormente, se diseña un modelo predictivo que podrá ser implementado a través de un sistema informático para la automatización de la generación de informes y visualización de resultados, lo que facilita la identificación de alumnos con posibles problemáticas de rendimiento académico. En la segunda etapa, se realizan procesos de intervención educativa con los estudiantes detectados con mayor posibilidad de reprobación de materias de acuerdo con los informes generados por el sistema predictivo y tomando en cuenta el análisis diagnóstico de los estudiantes sobre diversos aspectos que influyen en su aprendizaje.

#### **4.7 Población y Muestra**

El estudio se realiza con alumnos de nuevo ingreso pertenecientes a programas educativos de nivel superior de la Facultad de Matemáticas de la Universidad Autónoma de Yucatán, con sede en la ciudad de Mérida, Yucatán, México. En esta facultad operan seis programas de licenciatura: Ciencias de la Computación,

Ingeniería en Computación, Ingeniería de Software, Matemáticas, Enseñanza de las Matemáticas y Actuaría, sin embargo, se trabaja con las carreras del área de computación. En la primera etapa, se utilizan los datos disponibles de alumnos que han concluido el primer período escolar. En la segunda etapa, se trabaja con los alumnos del programa de estudio en donde se ha concentrado la mayor cantidad de estudiantes en riesgo escolar.

#### **4.8 Estrategia de Recolección y Análisis de Datos**

En la primera etapa, se utilizan los métodos y técnicas de clasificación de la MDE, así como las metodologías de desarrollo de sistemas. En el primer objetivo específico se busca diseñar modelos predictivos de rendimiento escolar mediante el uso de herramientas de minería de datos educativos, los datos disponibles se encuentran almacenados en varias fuentes de archivos digitales generados en los departamentos administrativos y académicos de la Institución, por lo que es necesario recuperarlos en su formato de origen, depurarlos y consolidarlos en un solo archivo en formato de hoja de cálculo Excel. Posteriormente, se realiza un procesamiento previo, requerido para aplicar técnicas de minería de datos educativas, utilizando el archivo Excel con la base de datos. El siguiente paso es exportar los datos a un archivo de texto CSV (valores separados por comas) o ARFF (formato de archivo de relación de atributos), para poder ser analizados mediante la herramienta de minería de datos WEKA.

En la segunda etapa se atiende el objetivo específico relacionado con la definición de estrategias de intervención educativa, para ello se diseñó un cuestionario diagnóstico que permite identificar elementos a considerar en la planificación de la intervención. Se recolectaron los datos aplicando el cuestionario a los alumnos de primer semestre para conocer su opinión respecto a las expectativas de calificaciones y materias reprobadas y contrastar con los resultados históricos. También para conocer su opinión sobre sus antecedentes, la percepción del nivel de exigencia, los servicios y apoyos que ofrece la escuela, las estrategias docentes,

estrategias con apoyo de tecnología educativa, asesorías y tutorías, entre otros aspectos. El cuestionario diagnóstico contiene preguntas tipo Likert y preguntas abiertas (Anexo 1). Se diseñó en la plataforma *Google Forms* para su aplicación vía Internet. De igual forma, las respuestas registradas en la plataforma fueron descargadas a un archivo de hoja de cálculo en formato Excel o CSV, para su análisis posterior.

En el caso del proceso de intervención, la validación de los resultados de la efectividad de las estrategias desplegadas es mediante técnicas estadísticas y el método de triangulación, al contrastar los datos de diversas fuentes, se utiliza el software SPSS. Se requiere, por tanto, el registro y análisis de datos provenientes de la aplicación de encuestas, entrevistas o fuentes alternas de datos, para ello, es necesaria la opinión de los participantes clave (alumnos, profesores, tutores o directivos). Como apoyo en el análisis general de los datos, en etapas cualitativas, se utiliza el software ATLAS.ti para agilizar el análisis y la elaboración de reportes. También es posible el uso de mapas conceptuales, diagramas causa-efecto, matrices, jerarquizaciones, organigramas o análisis de redes, según se requiera (Álvarez y Álvarez, 2014). Concluidas las etapas cuantitativas y cualitativas, se elabora y presenta un reporte de los hallazgos y resultados.

#### **4.9 Recursos Materiales y Humanos**

Se tiene la autorización de las autoridades de la escuela para utilizar los recursos disponibles en la misma y para obtener los datos requeridos por el proyecto. Los materiales de consulta están disponibles en la biblioteca del Campus, la cual pertenece al Consorcio Nacional de Recursos de Información Científica y Tecnológica, por lo que se tiene acceso a artículos y libros de gran variedad de fuentes científicas especializadas. En cuanto a los recursos tecnológicos, como servidores, redes, Internet, entre otros, también están disponibles en la Institución y el software empleado para el análisis de los datos es de libre distribución (WEKA, R, RStudio), y el software comercial fue adquirido o está disponible en la Institución

(Excel, SPSS, ATLAS.ti). En relación con la necesidad de aprendizaje de nuevas tecnologías informáticas o sistemas, se solventaron mediante el acceso a recursos audiovisuales, cursos en línea, libros y tutoriales disponibles a través de Internet.

Para el desarrollo y seguimiento de las actividades, hubo el interés por parte de la dirección de la Facultad de apoyar en la medida de lo posible las estrategias a implementar, en función de los recursos disponibles. El responsable de este proyecto de investigación, en coordinación con la responsable de la Secretaría Académica, planificaron reuniones y actividades, para que, junto con los profesores que imparten clases en los primeros semestres de la carrera y/o los responsables del programa de tutorías se diseñaran las líneas de acción necesarias para atender posibles problemas educativos identificados mediante el sistema y modelos desarrollados.

#### **4.10 Procedimientos**

En este apartado se presenta un resumen de los procedimientos seguidos durante el proyecto de investigación, que, por el enfoque mixto adoptado, incluye etapas y procesos con actividades particulares de acuerdo con las necesidades requeridas en cada momento específico y tratando de alcanzar los objetivos planteados. En la Figura 4, se muestran las etapas y actividades principales a realizar durante el trabajo de investigación. De acuerdo con la metodología de trabajo mixto en múltiples fases, los procesos pueden ejecutarse de manera simultánea o secuencial, dependiendo de la disponibilidad de los recursos, datos y tiempo o de las relaciones entre las etapas y procesos.



Figura 4 Diseño general del proyecto de investigación.

Se realizó la revisión de la literatura para construir el corpus del estado del arte y el marco teórico, partiendo de las preguntas y objetivos de la investigación. Para ello se definieron categorías y subcategorías de búsqueda y análisis, con base en ellas fueron realizadas búsquedas de información en repositorios científicos. Los resultados fueron filtrados por la relevancia para el estudio y luego analizados para identificar las principales características de los trabajos, alcances y vacíos que pudieran ser atendidos por nuevas investigaciones. Se realizó un informe en el cual se presenta el desarrollo temático y las conclusiones alcanzadas.

Durante el proceso de la Minería de Datos Educativos, se concluyeron las actividades que consistieron en recuperar los datos en crudo del entorno educativo; realizar tareas de depuración y limpieza de los datos; y una vez preparados los datos, se aplicaron técnicas de minerías de datos mediante el software. Se obtuvieron modelos con reglas para clasificar o predecir la posibilidad de riesgo académico de los estudiantes. Finalmente, se aplican técnicas de validación

cruzada para verificar la eficacia de las predicciones y se compararon con nuevos datos provenientes de alumnos que ingresaron en agosto de 2019.

Se realizó la implementación del sistema predictivo, a partir de los resultados y modelos generados en la etapa de la MDE. Para la programación se empleó el entorno de desarrollo de aplicaciones NetBeans 8.2, utilizando el lenguaje de programación Java. La decisión se tomó con base en la compatibilidad con el entorno de WEKA y la disponibilidad de interfaces de programación para acceder a datos en la plataforma Moodle, donde se encontraba la información de los alumnos. El sistema incluye los modelos con las reglas de clasificación para la predicción. Es necesario establecer mecanismos de conexión a la plataforma Moodle para el acceso a los datos de nuevas instancias de alumnos. Posteriormente se prueba el prototipo con datos de alumnos ficticios y se hacen los ajustes pertinentes para que, al comenzar el período escolar, pueda ser utilizado con los alumnos de nuevo ingreso en la detección de posibles casos de riesgo escolar.

Durante el proceso de investigación-acción, el primer paso fue diagnosticar las problemáticas implicadas en el rendimiento escolar, por lo que se aplicó una encuesta diagnóstica. En los resultados se detectaron algunas áreas de mejora en aspectos pedagógicos, de la operación de los planes de estudio y de los servicios. Se vio que es necesario atender situaciones de índole personal, apoyos para mejorar habilidades de aprendizaje y hábitos de estudio, especialmente es necesario apoyar a los estudiantes para que logren una pronta transición y adaptación a la universidad. Con la identificación de algunos de los factores que inciden en el rendimiento escolar, el siguiente paso fue, en conjunto con el grupo de profesores involucrados, planificar la implementación de estrategias de intervención educativa. Las necesidades más urgentes fueron:

- Implementar medidas de intervención educativa, especialmente para lograr una pronta transición y adaptación a la universidad.

- Mejorar y modificar programas de asesorías y tutorías, con alternativas de apoyo más eficaces y viables para los estudiantes.
- Atender aspectos de hábitos de estudio, temas académicos particulares, habilidades para el aprendizaje, situaciones personales y de antecedentes académicos.

La principal estrategia de intervención académica adoptada, y que atiende las necesidades más urgentes de la población estudiantil sujeta a la intervención, consiste en una modalidad de tutoría denominada tutoría de asignatura, en nuestro caso, basada en el diseño y desarrollo de un curso en línea en el que se revisan aspectos para la mejora del aprendizaje y adaptación a la escuela tratando de favorecer el desarrollo de procesos cognitivos y metacognitivos. Durante su impartición, y con apoyo del sistema desarrollado, se emplean métodos de la analítica del aprendizaje y mecanismos para evaluar el rendimiento y la satisfacción de varios aspectos relacionados con sus asignaturas y su desempeño personal, así como para obtener información útil para la toma de decisiones académicas. Se incluyeron medidas complementarias de apoyo en la búsqueda de la mejora del rendimiento académico de los estudiantes, mediante la intervención de tutores y directivos escolares.

Los resultados del proceso de intervención se determinaron con base en el análisis de los datos generados por encuestas y entrevistas a los participantes. La elaboración del reporte final incluye el análisis y contrastación de los resultados de todas las etapas y procesos, con el propósito de determinar el cumplimiento de los objetivos generales del estudio.

## 5. PROCESO DE MINERÍA DE DATOS EDUCATIVOS

El contenido de este capítulo describe las actividades realizadas en el marco del proceso de la minería de datos educativos. La Figura 5 ilustra las tareas principales de la MDE. Las etapas de extracción de datos y su procesamiento previo son fundamentales para contar con los insumos necesarios en las etapas posteriores, se describen las fuentes de los datos y las decisiones tomadas para contar con datos suficientes y confiables. En la etapa de aplicación de técnicas, se determinó el uso de varios algoritmos de clasificación, los resultados alcanzados se muestran en la sección de generación de modelos y validación de los modelos predictivos de riesgo académico. Finalmente, se realiza una interpretación de los modelos obtenidos, además, se presentan conclusiones preliminares de esta etapa del proyecto de investigación.



Figura 5 Tareas realizadas en la etapa de minería de datos educativos.

### 5.1 Extracción de Datos

El estudio fue desarrollado en la Facultad de Matemáticas de la Universidad Autónoma de Yucatán (UADY). Para propósitos de la investigación nos enfocamos exclusivamente en tres de los programas correspondientes al área de computación. Los programas de esta área son: Licenciatura en Ingeniería en Computación (LIC), Licenciatura en Ciencias de la Computación (LCC) y Licenciatura en Ingeniería de Software (LIS).

Para iniciar el proceso de generación de modelos predictivos, el primer paso consistió en ubicar los datos institucionales disponibles, los cuales pudieran aportar información relacionada con factores asociados al rendimiento académico de los

alumnos de interés para el estudio. Fue seleccionada información académica, sociodemográfica y de resultados de exámenes de ingreso de las fuentes indicadas a continuación:

- Las calificaciones de las asignaturas de los alumnos en su primer semestre de estudios se recuperaron de los registros del Departamento de Control Escolar. Desde 2016, para los programas LCC y LIS, puesto que es el año en que entraron en operación los nuevos planes de estudios basados en Modelo Educativo para la Formación Integral (MEFI) de la Universidad; el cual considera una calificación de 70 como la mínima aprobatoria, en cambio en el modelo anterior la calificación mínima era de 60. El plan de estudios de LIC ya se encontraba en operación desde 2013. Por tanto, los datos abarcan el período 2016 a 2019. Esta información fue proporcionada en 10 documentos de hojas de cálculo Excel, por lo que fue necesario homologar su estructura y consolidar los datos en un solo archivo.
- Los resultados del examen de ingreso Exani-II de los alumnos admitidos a los programas de estudios de las carreras del área de computación, entre 2016 y 2019. Esta información la proporcionó la Secretaría Académica mediante la entrega de 5 archivos en formato XLSX.
- Datos sociodemográficos de los alumnos de las carreras del área de computación. Esta información fue proporcionada por la Secretaría Académica a través de 4 archivos en formato XLSX.

Toda la información se recolectó en un momento único, cuando a la ocurrencia del fenómeno de estudio y los factores que influyeron en sus resultados ya habían sucedido. Al momento de consolidar la información, proveniente de todas las fuentes identificadas, se tuvo que hacer coincidir cada uno de los atributos o variables con el correspondiente alumno, mediante operaciones de ordenamiento, filtrado y copiado de datos. Finalmente, se obtuvo una sola base de datos almacenada en un archivo en formato Excel. El total de instancias obtenidas

corresponden a 415 alumnos, 90 de la Licenciatura en Ingeniería en Computación, 101 de la Licenciatura en Ciencias de la Computación y 224 de la Licenciatura en Ingeniería de Software. Cada uno de los registros de alumnos contaba con 65 atributos.

La confidencialidad de los datos de los estudiantes es un aspecto importante para el desarrollo del proyecto. En este sentido, los datos personales recuperados fueron manejados exclusivamente por el responsable del proyecto, quien también fue el único autorizado para mantener en resguardo las copias de los archivos de bases de datos. A partir de la obtención y consolidación de los datos en un solo archivo, también se eliminaron los atributos de nombres o matrículas escolares, con lo cual se evitó la identificación de los alumnos, de esta forma se garantiza el anonimato en la manipulación de la información en el proceso de análisis. Los modelos generados presentan información de forma genérica y sin mostrar datos de forma individual.

## **5.2 Procesamiento previo**

Una vez unificados los datos su preparación resultó compleja y tardada, sin embargo, es una tarea fundamental para poder hacer un análisis más preciso de la información. Se trabajó en el Excel empleando las funciones disponibles para la edición de los datos. De los 65 atributos disponibles, mediante un proceso de limpieza y eliminación de atributos no relevantes para el estudio, fueron retirados manualmente de la base de datos los campos cuya información fue identificada fácilmente como irrelevante o poco útil. También, 5 variables fueron agregadas, derivadas a partir de otros atributos, por ejemplo, a partir de la fecha de nacimiento se obtuvo la edad del estudiante, a partir de las calificaciones se obtuvo el promedio y la cantidad de materias aprobadas, del nombre de la escuela de procedencia se obtuvo la variable Prepa, para indicar si estudió en preparatorio perteneciente a la UADY o no.

En este paso se destaca la importancia y necesidad de convertir, normalizar o discretizar las variables para obtener mejores resultados o para poder utilizar algoritmos que requieren ciertas características en el formato de las variables. Los atributos de los alumnos elegido fueron aquellos que pudieran aportar información significativa para los propósitos del análisis. Un primer conjunto de variables se muestra en la Tabla 2.

Tabla 2 Descripción del conjunto de variables inicial.

<b>NOMBRE</b>	<b>DESCRIPCIÓN</b>
Carrera	Carrera que estudia
Sexo	Sexo del alumno
ICNE	Prueba competencias básicas (CB) del EXANI II
DIAG	Prueba de competencias disciplinares (CD) del EXANI II
Exani	Resultado global del examen de ingreso (CB 70% + CD 30%)
A_ing	Año de ingreso
N_insc	Número de materias inscritas
Edad	Edad al momento del ingreso
Mun	Municipio de procedencia
Edo	Estado de procedencia
Esc_Proc	Escuela de procedencia
Prepa	Preparatoria de procedencia (atributo derivado {UADY, No UADY})
E_Civil	Estado civil
N_Hijos	Número de hijos
S_Medico	Tipo de servicio médico
Resp	Responsable del alumno
Beca	Cuenta con beca
P_Aprob	Porcentaje de materias aprobadas
PROM	Promedio en primer semestre
RIESGO_A	Riesgo académico (P_Aprob < 0.8="SI", P_Aprob ≥ 0.8="NO")
RIESGO_P	Riesgo académico (PROM < 75="SI", PROM ≥ 75="NO")

Las variables dependientes RIESGO\_A y RIESGO\_P, se derivaron a partir de las variables PROM y P\_Aprob. Fue necesario definirlas como variables nominales o

categorías (clases), ya que varios de los algoritmos utilizados la clasificación y generación de los modelos predictivos requieren este tipo de dato en la definición de las clases.

En el proceso de asignación de valores a la variable RIESGO\_A, un valor de 0.8 o mayor del porcentaje de materias aprobadas (P\_Aprob), se asignó como “NO” a la variable RIESGO\_A, indicando que no hay riesgo académico, en caso contrario se asigna un “SI”. Dado que en primer semestre regularmente se inscriben 5 materias, un valor P\_Aprob menor a 0.8 indicaría que eventualmente reprobaría dos o más materias lo que se consideró como un riesgo académico alto para el alumno. De igual forma se procedió con la variable RIESGO\_P, en cuyo caso se consideró un valor de PROM mayor o igual a 75 para indicar que “NO” hay riesgo académico, en caso contrario se asignaría un “SI” a la variable.

Por lo tanto, quedó definido que un promedio de calificación inferior a 75 o un porcentaje de materias aprobadas inferior a 0.8 son indicadores de un bajo rendimiento académico, y en consecuencia existe alta probabilidad de riesgo académico para el estudiante. Existen estudios en los que se adoptaron criterios similares para la determinación de perfiles de rendimiento académico (Baker et al., 2015; Martínez et al., 2015; Padua, 2019; Río-Jenaro, Calle, Martín y Robaina, 2018).

Para poder analizar los datos en el software WEKA, fue necesario convertir el archivo de base de datos del formato Excel (XLSX) a un archivo de texto en formato CSV (valores separados por comas). Posteriormente, desde WEKA se leyó el archivo y se exportó al formato propio de esta herramienta, el cual utiliza archivos ARFF (formato de archivo de relación de atributos). De esta manera, todo quedó preparado para iniciar el proceso de análisis de la información.

### **5.3 Aplicación de Técnicas de MDE**

Durante esta etapa se empleó el software de minería de datos WEKA para realizar tareas de selección de atributos, también para la obtención validación de los modelos predictivos generados. Se utilizaron herramientas para la experimentación, y refinamiento de los parámetros, que permitieron generar modelos más precisos.

#### **5.3.1 Herramientas de la MDE**

Hay disponibles varias herramientas que hacen factible la aplicación de las técnicas de la MDE, las cuales incluyen algoritmos para realizar distintos análisis de datos educativos. Algunas de las más destacadas, de acuerdo a Slater et al. (2017), son: RapidMiner, DBminer, Waikato Environment for Knowledge Analysis (WEKA), KEEL, KoNstanz Information MinEr (KNIME), Orange y *Statistical Product and Service Solutions* (SPSS); además se han desarrollado paquetes o módulos para ser utilizados con Python y R.

Las herramientas anteriores cuentan con un amplio repertorio de algoritmos y técnicas, que las hacen apropiadas para la mayoría de las tareas de la MDE, incluyen los principales métodos que han sido utilizados en el análisis del rendimiento académico. También, como lo indica (Slater et al., 2017), hay herramientas diseñadas específicamente para otras tareas de la MDE, como la visualización o el análisis de redes sociales, entre otras, por ejemplo, el software *Graphical Interactive Student Monitoring* (GISMO)

En la revisión de la literatura, se señala el uso constante del software de minería de datos WEKA, el cual es un ambiente de simulación computacional que presenta un vasto soporte para la experimentación, con varios métodos estadísticos y de inteligencia artificial (Witten, Frank y Hall, 2011). El ser de libre distribución y de código abierto ha sido un factor importante para ganar popularidad en la comunidad de investigadores en el área de minería de datos. WEKA cuenta con un extenso repertorio de algoritmos de clasificación, predicción, agrupamiento y minería de asociaciones. También dispone de técnicas para aplicar varios criterios en la

selección de atributos y comparar su desempeño, de tal forma que se pueda elegir el subconjunto de variables que mejores resultados obtengan para la clasificación.

Además, incluye varios métodos de validación y verificación de la precisión de los resultados. Puede usarse en línea de comandos, a través de su interfaz gráfica o mediante la interfaz de programación del lenguaje Java. Los modelos generados pueden ser almacenados en archivos para, posteriormente, ser utilizados con nuevos datos y realizar predicciones. Una característica importante es la posibilidad de desarrollar aplicaciones Java e incluir las clases WEKA. Esto permite recuperar datos, crear instancias de métodos de clasificación a partir de modelos previamente generados en WEKA, y utilizarlas para evaluar nuevos casos, entre otras capacidades. Por lo anterior, se decidió la utilización del software WEKA versión 3.8.3, para el desarrollo del presente trabajo.

### 5.3.2 Selección de Variables Significativas

A partir del conjunto inicial de variables, se emplearon algunas de las técnicas integradas en el software WEKA para seleccionar un conjunto más reducido de variables significativas y que, posteriormente, se utilizan en el análisis y la generación de los modelos predictivos. Los resultados mostrados en la Tabla 3 corresponden a tres de los métodos de evaluación de atributos: CfsSubsetEval, WrapperSubsetEval con J48 y WrapperSubsetEval con NaiveBayes.

Tabla 3 Resultados de 3 métodos de selección de atributos.

Atributo	CfsSubsetEval/ GreedyStepwise	WrapperSubsetEval / J48/ BestFirst	WrapperSubsetEval/ NaiveBayes/ GreedyStepwise
Carrera *	0%	70%	80%
Sexo	0%	10%	50%
ICNE *	30%	0%	40%
DIAG *	100%	60%	0%
Exani *	100%	70%	80%
A_ing	0%	20%	50%
N_insc	0%	0%	20%
Edad *	0%	50%	40%
Mun *	80%	0%	10%
Edo	0%	20%	10%
Esc_Proc	20%	0%	10%

Prepa *	0%	30%	40%
E_Civil *	10%	40%	50%
N_Hijos	0%	10%	10%
S_Medico	0%	30%	10%
Resp *	0%	30%	60%
Beca	0%	10%	0%

El método CfsSubsetEval pondera de manera individual cada atributo, mientras que el método WrapperSubsetEval utiliza un algoritmo de clasificación para poder hacer la medición del rendimiento de los atributos que se están evaluando.

La Tabla 4 contiene los resultados del método CorrelationAttributeEval para la selección de atributos. En este caso del método CorrelationAttributeEval, clasifica y ordena los atributos con base en la correlación que tienen con la clase.

Tabla 4 Resultados del cuarto método de selección de atributos.

CorrelationAttributeEval / Ranker		
Mérito promedio	Rango promedio	Atributo
0.438 +- 0.014	1.2 +- 0.4	5 Exani *
0.43 +- 0.014	1.8 +- 0.4	4 DIAG *
0.407 +- 0.015	3 +- 0	3 ICNE *
0.247 +- 0.017	4 +- 0	12 Prepa *
0.172 +- 0.017	5 +- 0	8 Edad *
0.112 +- 0.009	6.5 +- 0.5	1 Carrera *
0.108 +- 0.015	6.9 +- 1.45	16 Resp *
0.088 +- 0.006	9 +- 0.63	11 Esc_Proc
0.088 +- 0.007	9.2 +- 1.54	7 N_insc
0.084 +- 0.01	9.8 +- 1.47	6 A_ing
0.082 +- 0.007	10 +- 0.89	9 Mun
0.058 +- 0.015	12.4 +- 1.28	15 S_Medico
0.06 +- 0.016	12.5 +- 0.92	10 Edo
0.032 +- 0.018	14.8 +- 1.4	13 E_Civil
0.023 +- 0.012	15.2 +- 0.98	17 Beca
0.02 +- 0.012	15.4 +- 1.11	2 Sexo
0.012 +- 0.008	16.3 +- 0.9	14 N_Hijos

Los resultados de los diferentes métodos no coincidieron totalmente, ya que cada uno captura de forma diferente la información y la relevancia de las variables. Se optó por elegir un subconjunto considerando los atributos en los que había un valor de relevancia regular o alto para dos o más métodos, procurando de esta manera no eliminar aún atributos que pudieran ser relevantes en las predicciones. Las

variables seleccionadas en esta etapa son las indicadas con (\*) en la Tabla 3 y la Tabla 4.

Por lo tanto, el subconjunto de variables independientes o predictoras quedó definido como: Carrera (programa educativo elegido), Exani (Examen Nacional de Ingreso, que muestra el resultado global ponderado ICNE+DIAG), ICNE (Índice Ceneval, que refleja el resultado del examen Exani-II Admisión), DIAG (resultado del examen Exani-II Diagnóstico, para ingenierías y tecnología), Edad (edad al momento del ingreso), Mun (municipio de procedencia), Prepa (escuela preparatoria de procedencia), E\_Civil (estado civil), Resp (responsable económico del estudiante). Las variables dependientes o clases son: RIESGO\_A (riesgo académico determinado por la relación de materias aprobadas) y RIESGO\_P (riesgo académico determinado por el promedio).

### ***5.3.3 Aplicación de Técnicas de Clasificación***

Las técnicas de minería de datos educativas se emplearon considerando el subconjunto de variables determinado previamente. Las técnicas y algoritmos de clasificación fueron utilizados en esta etapa. Por cada técnica utilizada y para cada variable dependiente se obtuvo un modelo predictivo y los resultados de la precisión y rendimiento del clasificador. En todos los casos se utilizó el conjunto de datos completo y el proceso para evaluar los resultados consistió en emplear validación cruzada con 10 iteraciones.

Varios de los algoritmos de clasificación disponibles en WEKA fueron utilizados. Después de varias pruebas se eligieron los que presentaron inicialmente un mejor rendimiento y precisión, los cuales fueron: J48, RandomForest, LMT, Logistic y MultilayerPerceptron. Los modelos resultantes y la evaluación de su precisión se presentan en la sección siguiente.

## **5.4 Generación de Modelos Predictivos**

Los modelos obtenidos para cada algoritmo incluyen dos casos que corresponden a cada variable clasificadora: RIESGO\_P, basada en el promedio de calificaciones y RIESGO\_A, basada en el porcentaje de materias aprobadas. Ambas clases pueden tener dos estados finales del estudiante: 1) el estudiante está en riesgo académico (valor de la clase = "SI") y 2) el estudiante no está en riesgo académico (valor de la clase = "NO"). A continuación, se describen los algoritmos empleados y los modelos generados con ellos.

### **5.4.1 Algoritmo J48**

Un árbol de decisión es un método de aprendizaje supervisado que construye un árbol a partir de un conjunto de datos históricos (datos de entrenamiento), encontrando patrones predictivos inductivamente. Cada rama del árbol es una posible explicación al problema de clasificación o predicción planteado (reglas de decisión o clases). Cada regla se interpreta leyendo la secuencia de nodos y sus condiciones a partir de la raíz y hasta llegar al nodo hoja (Peña-Ayala, 2014).

El algoritmo consiste en calcular una razón de ganancia para cada variable, basado en un valor de entropía o incertidumbre; a menor incertidumbre más información aporta el atributo para explicar la clasificación. Se seleccionan los mejores atributos predictores para ir creando nodos de forma descendente y recursiva; los nodos más próximos a la raíz son los mejores predictores, y a partir de ellos el proceso se repite para generar ramas en el árbol con nuevos nodos a partir de los datos restantes. Si no es posible encontrar más predictores el árbol deja de crecer. Puede utilizar atributos nominales o continuos, ya que los valores continuos los convierte automáticamente a intervalos discretos. La variable dependiente debe ser de tipo nominal. Esta técnica permite procesar gran cantidad de variables, y muchas de las variantes de árboles son completamente no-paramétricas, es decir, no se requiere una distribución de población con características específicas (Menacho, 2017).

El método J48 de clasificación estadística es la implementación en WEKA del algoritmo C.45 (Quinlan, 1993). Ha sido uno de los más utilizados por su sencillez al momento de interpretar las reglas predictivas de sus modelos. Se puede usar tanto para clasificar o identificar factores determinantes en una población, como para predecir futuras instancias. La Tabla 5 muestra las reglas de conocimiento obtenidas a partir del árbol J48 para la clase RIESGO\_P y su interpretación en lenguaje natural.

Tabla 5 Reglas de conocimiento derivadas del árbol de decisión J48 para la clase RIESGO\_P.

#	REGLA DE CONOCIMIENTO	INTERPRETACIÓN EN LENGUAJE NATURAL	OCURRENCIAS
1	DIAG <= 65 & DIAG <= 44: SI	Si el valor de DIAG es menor o igual que 44, entonces, SI hay riesgo académico.	(32.0)
2	DIAG <= 65 & DIAG > 44 & Carrera = LCC & Prepa = No UADY: SI	Si el valor de DIAG es mayor a 44 y menor o igual a 65 y Carrera es igual a LCC y Prepa es igual a No UADY, entonces, SI hay riesgo académico.	(32.0/5.0)
3	DIAG <= 65 & DIAG > 44 & Carrera = LCC & Prepa = UADY & Resp = MADRE: SI	Si el valor de DIAG es mayor a 44 y menor o igual a 65 y Carrera es igual a LCC y Prepa es igual a UADY y la responsable es la Madre, entonces SI hay riesgo académico.	(6.0/1.0)
4	DIAG <= 65 & DIAG > 44 & Carrera = LCC & Prepa = UADY & Resp = PADRE & ICNE <= 1180: SI	Si el valor de DIAG es mayor a 44 y menor o igual a 65 y Carrera es igual a LCC y Prepa es igual a UADY y el responsable es el Padre y el ICNE es menor o igual a 1180, entonces, SI hay riesgo académico.	(9.0/2.0)
5	DIAG <= 65 & DIAG > 44 & Carrera = LIS & Resp = MADRE & DIAG <= 53: SI	Si el valor de DIAG es mayor a 44 y menor o igual a 65 y Carrera es igual a LIS y la responsable es la Madre y DIAG es menor o igual a 53, entonces, SI hay riesgo académico.	(7.0)
6	DIAG <= 65 & DIAG > 44 & Carrera = LIS & Resp = MADRE & DIAG > 53 & Prepa = UADY & ICNE <=1216: SI	Si el valor de DIAG es mayor a 44 y menor o igual a 65 y Carrera es igual a LIS y la responsable es la Madre y DIAG es mayor a 53 y Prepa es UADY y el ICNE es menor o igual a 1216, entonces, SI hay riesgo académico.	(4.0)
7	DIAG <= 65 & DIAG > 44 & Carrera = LIS & Resp = PADRE & ICNE <= 1150: SI	Si el valor de DIAG es mayor a 44 y menor o igual a 65 y Carrera es igual a LIS y el responsable es el Padre y el ICNE es menor o igual a 1150, entonces, SI hay riesgo académico.	(17.0/3.0)
8	DIAG <= 65 & DIAG > 44 & Carrera = LIS & Resp = PADRE & ICNE > 1150 & Prepa = No UADY & ICNE <=1198: SI	Si el valor de DIAG es mayor a 44 y menor o igual a 65 y Carrera es igual a LIS y el responsable es el Padre y el ICNE es mayor a 1150, pero menor o igual a 1198, y Prepa es No UADY, entonces, SI hay riesgo académico.	(27.0/9.0)
9	DIAG <= 65 & DIAG > 44 & Carrera = LIS & Resp = PADRE & ICNE > 1150 & Prepa = UADY & ICNE > 1174: SI	Si el valor de DIAG es mayor a 44 y menor o igual a 65 y Carrera es igual a LIS y el responsable es el Padre y el ICNE es mayor a 1174 y Prepa es UADY, entonces, SI hay riesgo académico.	(18.0/7.0)
10	DIAG <= 65 & DIAG > 44 & Carrera = LIS & Resp = UNO MISMO: SI	Si el valor de DIAG es mayor a 44 y menor o igual a 65 y Carrera es igual a LIS y el responsable es Uno mismo, entonces, SI hay riesgo académico.	SI (8.0/1.0)

Por la dimensión del árbol de decisión y para el propósito de identificar condiciones de riesgo académico, se presentan las reglas que indican los casos en que la variable dependiente tomará un valor de “SI”, es decir, se predice la existencia de riesgo académico. Lo que significa que los alumnos podrían obtener un promedio de calificación inferior a 75 en el semestre.

El atributo DIAG es el más significativo, ya que a partir de él se derivaron todas las reglas de clasificación o predicción listadas en la Tabla 5. La columna de ocurrencias muestra la cantidad de casos que se presentaron para cada regla, tanto aciertos como errores en la clasificación, lo cual puede servir como indicador para detectar los atributos y condiciones más frecuentes que determinan situaciones de riesgo académico de los alumnos. De forma similar, en la Tabla 6 se muestran las reglas de conocimiento obtenidas a partir del árbol J48 para la clase RIESGO\_A y su interpretación en lenguaje natural.

Tabla 6 Reglas de conocimiento derivadas del árbol de decisión J48 para la clase RIESGO\_A.

#	REGLA DE CONOCIMIENTO	INTERPRETACIÓN EN LENGUAJE NATURAL	OCURENCIAS
1	ICNE <= 1216 & DIAG <= 62 & Carrera = LIC & Prepa = No UADY: SI	Si el valor de ICNE es menor o igual que 1216 y DIAG es menor o igual que 62 y la Carrera es igual a LIC y Prepa es igual a No UADY, entonces, SI hay riesgo académico.	(40.0/15.0)
2	ICNE <= 1216 & DIAG <= 62 & Carrera = LIC & Prepa = UADY & Edad > 21: SI	Si el valor de ICNE es menor o igual que 1216 y DIAG es menor o igual que 62 y la Carrera es igual a LIC y Prepa es igual a UADY y Edad es mayor a 21, entonces, SI hay riesgo académico.	(5.0)
3	ICNE <= 1216 & DIAG <= 62 & Carrera = LCC & Prepa = No UADY: SI	Si el valor de ICNE es menor o igual que 1216 y DIAG es menor o igual que 62 y la Carrera es igual a LCC y Prepa es igual a No UADY, entonces, SI hay riesgo académico.	(49.0/6.0)
4	ICNE <= 1216 & DIAG <= 62 & Carrera = LCC & Prepa = UADY & Exani <= 781.25: SI	Si el valor de ICNE es menor o igual que 1216 y DIAG es menor o igual que 62 y la Carrera es igual a LCC y Prepa es igual a UADY y Exani es menor o igual a 781.25, entonces, SI hay riesgo académico.	(11.0/1.0)
5	ICNE <= 1216 & DIAG <= 62 & Carrera = LIS: SI	Si el valor de ICNE es menor o igual que 1216 y DIAG es menor o igual que 62 y la Carrera es igual a LIS, entonces, SI hay riesgo académico.	(88.0/25.0)
6	ICNE <= 1216 & DIAG > 62 & Prepa = No UADY & Edad <= 23 & Exani <= 803.25: SI	Si el valor de ICNE es menor o igual que 1216 y DIAG es mayor 62 y la Prepa es igual a No UADY y Edad es menor o igual a 23 y Exani es menor o igual a 803.25, entonces, SI hay riesgo académico.	(10.0/2.0)
7	ICNE <= 1216 & DIAG > 62 & Prepa = No UADY & Edad <= 23 & Exani > 803.25 & ICNE > 1192 & Edad <=20 & Edad > 18: SI	Si el valor de ICNE está entre 1193 y 1216 y DIAG es mayor 62 y la Prepa es igual a No UADY y Exani es mayor a 803.25 y Edad está entre 18 y 20 años, entonces, SI hay riesgo académico.	(6.0/1.0)
8	ICNE <= 1216 & DIAG > 62 & Prepa = No UADY & Edad > 23: SI	Si el valor de ICNE es menor o igual que 1216 y DIAG es mayor 62 y la Prepa es igual a No UADY y Edad mayor a 23, SI hay riesgo académico.	(4.0)
9	ICNE > 1216 & Edad > 25: SI	Si el valor de ICNE es mayor a 1216 y Edad es mayor a 25, SI hay riesgo académico.	(4.0)

Para este caso, se observa que el atributo ICNE tiene el mayor peso predictivo, pues es el punto de partida para todas las reglas obtenidas. Hay en total nueve reglas de decisión que devuelven un resultado igual a “SI”, indicando el riesgo de reprobar dos o más asignaturas en el semestre.

#### **5.4.2 Algoritmo RandomForest**

Los bosques aleatorios son una combinación de predictores de árboles de decisión. El método genera múltiples versiones del árbol tomando muestras aleatorias de los datos de entrenamiento y promediando estas para obtener un mejor clasificador (Breiman, 2001). Es posible que se genere una gran cantidad de árboles, incluso similares. Por ello, el algoritmo escoge un subconjunto de variables predictivas antes de dividir cada nodo, para obtener árboles diferentes no correlacionados del conjunto de datos. Para construir el modelo final, cada árbol individual en el bosque aleatorio proporciona una predicción de clase y la clase con más votos se convierte en la predicción del modelo.

El modelo resultante fue generado utilizando 100 iteraciones con parámetros de aprendizaje por defecto. La visualización del bosque de árboles no es posible debido a la gran cantidad de árboles obtenidos, sin embargo, WEKA tiene la opción de salvar en archivo los modelos, lo cual se realizó para los resultados conseguidos con las dos variables de clase RIESGO\_A y RIESGO\_P. Los archivos con los modelos pueden ser cargados en WEKA o ser utilizados desde programas computacionales, que incluyan las librerías apropiadas, para realizar la predicción de nuevos casos de alumnos.

#### **5.4.3 Algoritmo LMT**

Este algoritmo construye clasificadores de árboles de modelo logístico (LMT, por sus siglas en inglés), que son árboles de clasificación con funciones de regresión logística en las hojas. El resultado de combinar modelos de estructuras de árboles y la regresión logística produce un árbol con estimaciones de probabilidad de clase explícitas en lugar de sólo clasificar (Landwehr, Hall y Frank, 2006).

LMT produce un solo árbol con divisiones binarias en atributos numéricos, divisiones de múltiples vías en variables nominales y modelos de regresión logística en las hojas. A pesar de no ser tan sencilla su interpretación, es más entendible que los resultados producidos por otros métodos, como en el caso de modelos con múltiples árboles. También el algoritmo se asegura de dejar sólo los atributos más relevantes en el modelo.

En la Figura 6 se muestra el árbol LMT. El árbol tiene sólo dos nodos hoja o terminales, correspondientes a los dos posibles valores de la clase RIESGO\_A. Cada valor de la clase se modela como una función de variables explicativas que determinan la probabilidad de un posible resultado.

```

Class SI :
3.49 +
[Carrera=LIC] * -0.55 +
[DIAG] * -0.02 +
[Exani] * -0 +
[Edad] * 0.06 +
[Prepa=UADY] * -0.38 +
[Resp=PADRE] * -0.13

Class NO :
-3.49 +
[Carrera=LIC] * 0.55 +
[DIAG] * 0.02 +
[Exani] * 0 +
[Edad] * -0.06 +
[Prepa=UADY] * 0.38 +
[Resp=PADRE] * 0.13

```

Figura 6 Árbol de modelo logístico (LMT) para la clase RIESGO\_A.

La Figura 7 muestra el árbol LMT resultante para la clase RIESGO\_P. Se observan similitudes en entre los dos árboles, al considerar los mismos atributos como factores relevantes en la predicción, variando en los coeficientes de las funciones logísticas de los nodos terminales.

```

Class SI :
3.99 +
[Carrera=LIC] * -0.77 +
[DIAG] * -0.03 +
[Exani] * -0 +
[Edad] * 0.07 +
[Prepa=UADY] * -0.22 +

```

```

[Resp=MADRE] * 0.08 +
[Resp=UNO MISMO] * 0.4 +
[Resp=HERMANO(A)] * -0.66 +
[Resp=OTRO] * 0.56

Class NO :
-3.99 +
[Carrera=LIC] * 0.77 +
[DIAG] * 0.03 +
[Exani] * 0 +
[Edad] * -0.07 +
[Prepa=UADY] * 0.22 +
[Resp=MADRE] * -0.08 +
[Resp=UNO MISMO] * -0.4 +
[Resp=HERMANO(A)] * 0.66 +
[Resp=OTRO] * -0.56

```

Figura 7 Árbol de modelo logístico (LMT) para de la clase RIESGO\_P.

El cálculo de la predicción se realiza sumando los elementos de la función. Cuando hay algún atributo nominal, por ejemplo, Carrera=LIC, en caso de que sea verdadera la condicional se multiplica 1 por el valor indicado del peso de la variable; en caso de ser un atributo numérico se multiplica directamente por el valor del peso de la variable. Si el resultado de la función es 0, significa que la clase toma el valor SI, que en nuestro modelo equivale a un resultado de riesgo académico.

#### 5.4.4 Algoritmo Logistic

Los modelos de regresión logística lineal permiten predecir el resultado de una variable dependiente categórica. La regresión logística es un instrumento estadístico de análisis multivariado. Su propósito es predecir la probabilidad de que ocurra un evento (Le Cessie y Van Houwelingen, 1992).

En la Figura 8 se muestra el modelo logístico para la clase o variable dependiente RIESGO\_A.

Coefficients...	
Variable	Class
	SI
Carrera=LIC	-0.915
Carrera=LCC	0.2996
Carrera=LIS	0.4034
ICNE	-0.0046
DIAG	-0.0699

<i>Exani</i>	-0.0044
<i>Edad</i>	0.141
<i>Prepa=UADY</i>	-0.5613
<i>Resp=MADRE</i>	0.3755
<i>Resp=PADRE</i>	0.2016
<i>Resp=UNO MISMO</i>	0.8621
<i>Resp=HERMANO(A)</i>	0.4544
<i>Resp=TIO(A)</i>	-74.7232
<i>Resp=OTRO</i>	1.3282
<i>Intercept</i>	10.1386
<i>Odds Ratios...</i>	
<i>Variable</i>	<i>Class</i>
	<i>SI</i>
=====	
<i>Carrera=LIC</i>	0.4005
<i>Carrera=LCC</i>	1.3494
<i>Carrera=LIS</i>	1.497
<i>ICNE</i>	0.9954
<i>DIAG</i>	0.9325
<i>Exani</i>	0.9956
<i>Edad</i>	1.1514
<i>Prepa=UADY</i>	0.5705
<i>Resp=MADRE</i>	1.4558
<i>Resp=PADRE</i>	1.2234
<i>Resp=UNO MISMO</i>	2.3682
<i>Resp=HERMANO(A)</i>	1.5752
<i>Resp=TIO(A)</i>	0
<i>Resp=OTRO</i>	3.7742

Figura 8 Coeficientes del modelo logístico para la clase RIESGO\_A.

De igual forma, la Figura 9 muestra las ponderaciones de los atributos en el modelo logístico para la clase RIESGO\_P.

<i>Coefficients...</i>	
<i>Variable</i>	<i>Class</i>
	<i>SI</i>
=====	
<i>Carrera=LIC</i>	-1.1655
<i>Carrera=LCC</i>	0.4602
<i>Carrera=LIS</i>	0.4556
<i>ICNE</i>	-0.005
<i>DIAG</i>	-0.0777
<i>Exani</i>	-0.0049
<i>Edad</i>	0.1472
<i>Prepa=UADY</i>	-0.4098
<i>Resp=MADRE</i>	0.4287
<i>Resp=PADRE</i>	0.2634
<i>Resp=UNO MISMO</i>	1.1287
<i>Resp=HERMANO(A)</i>	-1.4086
<i>Resp=TIO(A)</i>	-78.285
<i>Resp=OTRO</i>	1.6384
<i>Intercept</i>	11.135
<i>Odds Ratios...</i>	
	<i>Class</i>

<i>Variable</i>	<i>SI</i>
<i>Carrera=LIC</i>	<i>0.3118</i>
<i>Carrera=LCC</i>	<i>1.5845</i>
<i>Carrera=LIS</i>	<i>1.5772</i>
<i>ICNE</i>	<i>0.995</i>
<i>DIAG</i>	<i>0.9253</i>
<i>Exani</i>	<i>0.9951</i>
<i>Edad</i>	<i>1.1586</i>
<i>Prepa=UADY</i>	<i>0.6638</i>
<i>Resp=MADRE</i>	<i>1.5353</i>
<i>Resp=PADRE</i>	<i>1.3014</i>
<i>Resp=UNO MISMO</i>	<i>3.0916</i>
<i>Resp=HERMANO(A)</i>	<i>0.2445</i>
<i>Resp=TIO(A)</i>	<i>0</i>
<i>Resp=OTRO</i>	<i>5.1468</i>

Figura 9 Coeficientes del modelo logístico para la clase RIESGO\_P.

Los coeficientes son las ponderaciones que se aplican a cada atributo antes de sumarlos. El resultado de la suma es la probabilidad de que la nueva instancia pertenezca a la clase. En los modelos presentados, un valor mayor 0.5 significa que las variables dependientes tendrán un valor igual a “SI”, lo que indica una predicción positiva de riesgo académico. Las razones de probabilidades (*odds ratios*) indican qué tanta influencia o efecto tendrá un cambio de ese valor en la predicción.

#### **5.4.5 Algoritmo MultilayerPerceptron**

Se generó un modelo basado en red neuronal con propagación hacia atrás conocido como MultilayerPerceptron (MP). Este tipo de algoritmo emplea aprendizaje supervisado con el conjunto de entrenamiento y construye la red neuronal en una serie de niveles o capas, que determinan la clasificación de instancias. Deben existir al menos tres capas que son: capa de entrada, capa de procesamiento u oculta, y capa de salida. En cada capa puede existir un número variable de nodos o neuronas interconectadas que ejecutan funciones no lineales comúnmente de tipo sigmoideal o logística, en las que sea cual sea la entrada la salida estará comprendida entre 0 y 1. Cuando una neurona recibe una entrada, se activa y genera una salida que transmite a la siguiente capa, el valor transmitido dependerá del peso asignado a cada conexión (Mitra y Pal, 1995).

Los modelos para las clases RIESGO\_A y RIESGO\_P, tienen 10 nodos cada una, con los pesos de predicción que les fueron asignados en el proceso de aprendizaje. Un inconveniente de las redes neuronales es la dificultad de comprender los modelos que generan. En la Figura 10 se muestran 3 de los nodos del modelo MP.

```

Sigmoid Node 0
Inputs  Weights
Threshold -1.3064491512227503
Node 2   6.096383309821736
Node 3   -6.219122630052431
Node 4   -5.484905886307731
Node 5    1.839999062521567
Node 6    1.6192370537262368
Node 7    3.8194015616049812
Node 8   -2.921985664613799
Node 9    1.8538491040214233
Sigmoid Node 1
Inputs  Weights
Threshold  1.3064565421745287
Node 2   -6.096388571581784
Node 3    6.219104407340171
Node 4    5.484672734189203
Node 5   -1.8400167841509871
Node 6   -1.6198128315996037
Node 7   -3.8189497476356724
Node 8    2.9224180550225483
Node 9   -1.8539339831271278
Sigmoid Node 2
Inputs  Weights
Threshold  0.31547768643364943
Attrib Carrera=LIC  4.483306276354068
Attrib Carrera=LCC -2.7535846098268304
Attrib Carrera=LIS -2.0121033202057395
Attrib ICNE  0.18408561410797677
Attrib DIAG -5.689743762265199
Attrib Exani -2.1842360451003153
Attrib Edad  9.322224538137549
Attrib Prepa=UADY -7.526943203504304
Attrib Resp=MADRE -0.7600297549616539
Attrib Resp=PADRE  0.8530293486362223
Attrib Resp=UNO MISMO -5.12567774369935
Attrib Resp=HERMANO(A)  1.7827187040088694
Attrib Resp=TIO(A) -0.4131480848228954
Attrib Resp=OTRO  2.3742931048318376
...

```

Figura 10 Modelo parcial MultilayerPerceptron para la clase RIESGO\_P.

## 5.5 Validación del Modelo

En la Tabla 7 se muestra el resumen de los resultados obtenidos al calcular la precisión de los modelos. Se incluyen los valores de precisión y el área bajo la curva ROC para determinar la calidad de los clasificadores. Se empleó el método de validación cruzada con 10 iteraciones, utilizando las 415 instancias disponibles en la base de datos.

Tabla 7 Resultados de la evaluación de los modelos predictivos.

ALGORITMO	RIESGO_A		RIESGO_P	
	Precisión	AUC ROC	Precisión	AUC ROC
J48	69.8795 %	0.728	71.5663 %	0.736
RandomForest	68.9157 %	0.770	73.012 %	0.796
<b>LMT</b>	<b>71.0843 %</b>	<b>0.782</b>	<b>75.4217 %</b>	<b>0.805</b>
<b>Logistic</b>	<b>70.8434 %</b>	<b>0.780</b>	<b>74.9398 %</b>	<b>0.807</b>
MultilayerPerceptron	65.5422 %	0.721	74.6988 %	0.765

Al comparar los resultados generales entre la clase RIESGO\_A y RIESGO\_P se observa que la variable dependiente RIESGO\_P, definida a partir del promedio de calificaciones, tiene mejores valores en el indicador AUC ROC y también en los niveles de precisión. En la comparación de los valores de precisión y calidad de cada uno de los algoritmos, los mejores resultados fueron obtenidos por los modelos LMT y Logistic; interpretando los valores AUC ROC de ambos modelos, y siguiendo las pautas propuestas por (Minguillón et al., 2017), se puede decir que son buenos clasificadores y es factible su utilización para pronosticar el riesgo académico.

## 5.6 Interpretación de los Modelos Resultantes

En este apartado se comentan los resultados del proceso de selección de variables y algunas de las características de los modelos generados a partir de la aplicación

de varias de las técnicas de clasificación/predicciones disponibles en el software WEKA.

De acuerdo con los valores de relevancia de los atributos disponibles (Tabla 3 y Tabla 4), el mejor método de selección de variables, para el conjunto de datos en estudio, es *CorrelationAttributeEval/Ranker*, ya que 7 de los 9 atributos seleccionados coinciden con sus resultados. Haciendo caso a los méritos calculados para los tributos, las mejores variables predictoras son: *ICNE*, *DIAG*, *Exani* y *Carrera*, las cuales tienen el mayor potencial predictivo, lo que se ve reflejado en los modelos obtenidos.

En cuanto a los modelos predictivos, las reglas generadas por el árbol de decisión empleando el algoritmo J48, aportan mucha luz al entendimiento de los factores que influyen en el rendimiento académico, en especial para el propósito de este trabajo, ya permiten identificar con mucha claridad la serie de condiciones en las cuales los alumnos adquieren una condición de riesgo académico.

Considerando las reglas de conocimiento obtenidas por el algoritmo J48 para la clase RIESGO\_P, vistas en la Tabla 5, el atributo DIAG tiene el mayor peso como variable predictora. En el siguiente nivel, analizando las reglas con más ocurrencias, se observa, de forma general, que, si el atributo Carrera es LCC o LIS, y la preparatoria de origen es igual a No UADY, entonces hay posible riesgo académico. También, hay condiciones en que el responsable económico o el ICNE pueden contribuir como factores de riesgo académico.

Las reglas de decisión presentadas en la Tabla 6, fueron generadas por algoritmo J48, teniendo a la variable RIESGO\_A como su clase. Se observa que los atributos ICNE y DIAG aparecen como los atributos de mayor relevancia como variables predictoras. En los casos con mayores ocurrencias, la variable Carrera es factor de riesgo cuando los alumnos son del programa LIS; o si son LIC o LCC y su preparatoria de origen es No UADY. Por otro lado, se detectan condiciones en los

que a pesar de tener resultados aceptables en DIAG, Exani e ICNE, hay riesgo si Prepa es No UADY y la Edad del alumno es mayor a 18 años.

En general, en ambos árboles de decisión los resultados de las pruebas contenidas en el examen de ingreso Exani-II, tienen la mayor correlación respecto al rendimiento académico de los alumnos, ya sea que se considere la variable dependiente basada en el promedio de calificaciones o la basada en la cantidad de materias reprobadas. Hay coincidencia con los hallazgos de (Miguéis et al., 2018), quienes determinaron que los atributos con mayores pesos predictivos fueron promedios de calificaciones en exámenes de ingreso y calificaciones al primer año.

En menor medida se observan otros factores como: la carrera elegida, lo que nos habla de diferencias en la conformación de los grupos pertenecientes a los diferentes programas; la preparatoria de origen, indicando una posible dificultad de los alumnos de otras escuelas para adaptarse al modelo empleado en la UADY; el responsable del alumno y su edad, que posiblemente estén relacionados con factores de su entorno familiar o problemáticas socioeconómicas.

Estos hallazgos son consistentes con el trabajo presentado por (Montes y Lerner, 2010), en donde se comenta que el rendimiento académico no sólo se explica por las calificaciones obtenidas, sino que existen otros aspectos relacionados con la dimensión académica, la económica, la familiar, la personal y la institucional, que aportan a su comprensión; como también observaron (Merchan y Duarte, 2016).

Respecto a los modelos LMT mostrados en la Figura 6 y Figura 7, los atributos más relevantes son Carrera, DIAG, Exani, Edad, Prepa y Resp. Sin embargo, no es tan clara la interpretación del modelo como en el caso de los árboles de decisión, ya que el peso de los atributos en la función logística de los nodos terminales depende de si se trata de una variable nominal o numérica, y cuando es una variable numérica esta puede ser muy variada en el rango de sus valores, comparada contra otras variables numéricas del modelo. A pesar de lo anterior, estos modelos

obtuvieron la mejor precisión de todos los algoritmos evaluados, por lo que se recomienda su empleo en la determinación del riesgo académico de los alumnos.

El modelo generado con el algoritmo Logistic, resulta muy similar al modelo LMT, ya que genera una función de regresión logística con la que se calcula la probabilidad de ocurrencia de un evento o valor de la clase. Los atributos relevantes son *Carrera, ICNE, DIAG, Exani, Edad, Prepa y Resp*. Este modelo es el segundo mejor en cuanto a confiabilidad y resultados de predicción. Para este conjunto de variables las funciones logísticas parecen tener un buen desempeño.

El modelo RandomForest no se visualiza en los resultados que genera WEKA, debido a la gran cantidad de árboles obtenidos. Mientras que en el modelo MultilayerPerceptron (Figura 10) hay 10 nodos o neuronas, que contienen atributos y sus pesos, los cuales determinan el valor de la función sigmoïdal. Se observa que las variables empleadas son las mismas que en el modelo Logistic. La interpretación del modelo es difícil pues habría que dar seguimiento a las conexiones entre capas y nodos de la red neuronal.

Al comparar los atributos y clases utilizados en otras investigaciones (Berens et al., 2019; Buenaño-Fernández et al., 2019; Imran et al., 2019; Rico y Sánchez, 2018) se observa que las variables utilizadas corresponden principalmente a calificaciones de cursos, y las predicciones van en el sentido de aprobar sólo un curso, además, los datos requeridos deben ser obtenidos cuando el curso ya ha sido iniciado, en ocasiones con poco margen para desplegar estrategias de intervención educativa. Por el contrario, en nuestra propuesta los datos requeridos son accesibles previo el comienzo del semestre y los resultados de la predicción resultan oportunos para ajustar el diseño de los cursos, informar a las instancias escolares responsables de la atención y apoyo estudiantil o desplegar estrategias adicionales de intervención.

Analizando la precisión de los mejores algoritmos e interpretando el indicador AUC ROC, podemos decir que los resultados alcanzados son suficientemente buenos

para respaldar la pertinencia del empleo de los modelos en la predicción del riesgo académico de nuevas instancias de alumnos.

### **5.7 Conclusiones de la Etapa de MDE**

En esta etapa se experimentó con varias técnicas de minería de datos educativas para generar modelos predictivos de rendimiento académico con un buen nivel de confiabilidad y exactitud, que permitiera identificar casos de alumnos en situación de riesgo académico, pertenecientes a grupos de nuevo ingreso de las carreras del área de computación de la Facultad de Matemáticas de la Universidad Autónoma de Yucatán.

Como se observó en la literatura consultada, la determinación de rendimiento académico es un problema multifactorial en donde confluyen una gran cantidad de variables, no sólo de orden académico o del entorno escolar, sino del contexto sociodemográfico o de aspectos cognitivos e interpersonales, entre otros.

La tarea de recopilación y procesamiento de datos resultó crucial para obtener información de calidad y atributos relevantes para ser utilizados en las tareas de la minería de datos educativas. Se analizaron diversas fuentes de datos y como resultado se obtuvo un archivo con 415 instancias de alumnos cada uno con 65 atributos. El método empleado para seleccionar variables permitió reducir a 9 la cantidad de atributos significativos asociados a los alumnos en estudio, con el consiguiente ahorro de espacio de almacenamiento, reducción de complejidad y la mejora en la calidad de los modelos generados.

La aplicación de los métodos de clasificación se realizó mediante el software WEKA y se utilizó la base de datos acotada al subconjunto de variables elegidas. También se consideraron dos casos de variables dependientes (clases) para determinar valores de predicción. La primera tomando el porcentaje de materias reprobadas, en el que la reprobación de dos o más materias equivaldría a un valor afirmativo de riesgo (variable RIESGO\_A). En la segunda opción, se determinó una predicción de

riesgo académico cuando el promedio de calificaciones era inferior a 75 (variable RIESGO\_P).

Se generaron modelos predictivos mediante 5 algoritmos de clasificación de la minería de datos educativa. Con cada uno de ellos se emplearon las dos opciones de variables de clase y los valores de exactitud y predicción se contrastaron. De acuerdo con los resultados generales, la clase RIESGO\_P obtuvo en todos los algoritmos los mejores valores de exactitud y confiabilidad. De igual forma el mejor algoritmo predictivo resultó LMT, con un nivel de instancias correctamente clasificadas de 71.08%, para la variable RIESGO\_A y de 75.42% para la variable RIESGO\_P y sus valores del área bajo la curva ROC mostraron un buen nivel de confiabilidad, con valores 0.782 y 0.805 respectivamente.

De los atributos seleccionados, en general se observa que el resultado de la prueba de competencias disciplinares DIAG y la prueba diagnóstica de competencias básicas ICNE tienen el mayor poder predictivo y determinan en buena medida el bajo rendimiento escolar o riesgo académico; le siguen en importancia: el resultado final del Exani-II, la carrera elegida, la preparatoria de origen, la edad y el responsable del alumno.

También se pudo ver la combinación de factores que delinear a aquellos alumnos con tendencia a reprobar más asignaturas u obtener promedios de calificaciones más bajos. En general, cuando los alumnos de nuevo ingreso obtienen un puntaje bajo en DIAG o ICNE y la preparatoria de origen no pertenece a la UADY presentan riesgo académico; o en algunos casos, aun cuando tienen puntajes regulares en las pruebas que conforman el Exani-II, pero provienen de preparatoria no UADY o dependen de alguien diferente a su padre o su edad es mayor a 18 años, también se configura un escenario de riesgo académico.

Se ha mostrado que las diversas técnicas de la MDE pueden emplearse de manera muy flexible y práctica para seleccionar subconjuntos de atributos significativos y con ellos aplicar algoritmos de clasificación. Los resultados en los niveles de

precisión y confiabilidad de los algoritmos nos indican la posibilidad de emplear los modelos de manera efectiva, con el propósito de predecir el nivel de riesgo académicos de los estudiantes de nuevo ingreso los programas del área de computación.

El conocimiento adquirido con los modelos predictivos indica, por el lado académico, la necesidad de atender aspectos relacionados con los antecedentes académicos de los alumnos y tomar medidas para asegurar el entendimiento del modelo educativo de la UADY por todos los alumnos; pero, por otro lado, obliga a prestar atención en las condiciones socioeconómicas de los estudiantes, y considerar el despliegue de estrategias más efectivas de seguimiento y apoyo.

En trabajos futuros se continuará con la recopilación de datos para mejorar la precisión de los modelos predictivos, y se dará seguimiento a la implementación de los modelos mediante el desarrollo de un sistema informático, el cual será utilizado con las nuevas generaciones de estudiantes del área de computación para detectar de manera oportuna a estudiantes en situación de riesgo académico, lo cual dará la oportunidad de realizar intervenciones educativas que coadyuven a disminuir la problemática del bajo rendimiento académico.

## **6. PROCESO DE IMPLEMENTACIÓN DEL SISTEMA PREDICTIVO**

En este capítulo se describen las acciones realizadas en el proceso de desarrollo del sistema predictivo. Se presenta la arquitectura y los elementos tecnológicos empleados en su implementación, así como las principales funcionalidades de la aplicación. De igual forma se explica como los modelos generados mediante las técnicas de MDE se cargan desde el sistema para que sea posible realizar predicciones a partir de datos de alumnos de nuevo ingreso, dicha información fue recuperada de la plataforma Moodle, a partir de encuestas que permitían registrar las principales variables requeridas para realizar la predicción de riesgo y analizar el nivel de satisfacción de los alumnos respecto a sus cursos. Posteriormente, se explican las características de los módulos de predicción y de seguimiento, y cómo los resultados obtenidos se comunican a los protagonistas del proceso educativo. Se cierra el capítulo con algunos comentarios referentes al uso que se le puede dar al sistema y sus implicaciones.

### **6.1 Componentes y Tecnologías Utilizadas**

El desarrollo del sistema implicó una valoración de las tecnologías disponibles. Los modelos predictivos fueron generados mediante el software WEKA, y almacenados en archivos en un formato propio de esta plataforma. Para su utilización en un sistema desarrollado a la medida, la opción más viable era el empleo del lenguaje de programación Java, ya que los archivos de las clases con los algoritmos de MD y los modelos WEKA obtenidos pueden ser importados directamente mediante las librerías adecuadas. De igual forma, el acceso a la información desde la plataforma Moodle fue posible por la disponibilidad de una interfaz de programación de aplicaciones (API) que permite la recuperación de datos de los cursos y sus elementos, por ejemplo, calificaciones, tareas, cuestionarios, entre otros. También es posible el envío de datos desde el sistema hacia la plataforma Moodle, por ejemplo, para mandar mensajes, modificar calificaciones, etc.

Las tecnologías y elementos utilizados se enlistan a continuación:

- *IDE NetBeans 8.2*, entorno de desarrollo integrado para la edición, pruebas y generación de la aplicación. Cuenta con interfaces para la edición de ventanas y formularios, y la creación de las clases del proyecto.
- *Lenguaje Java*, para la programación y diseño del sistema, caracterizado por su enfoque orientado a objetos.
- *weka.jar*, archivo que contiene todas las clases necesarias para la manipulación de elementos WEKA, a través de la interfaz de programación.
- *org.json.jar*, archivo que contiene clases para acceder a objetos de la plataforma Moodle, a través del servicio de la API REST (*Representational State Transfer*), y una interfaz de intercambio de datos HTTP (*Hypertext Transfer Protocol*) basada en el formato de mensajes JSON (*JavaScript Object Notation*).
- *jfreechart-1.5.0.jar*, archivos con clases para generar de forma dinámica diferentes tipos de gráficas.

## **6.2 Arquitectura del Sistema**

En esta sección se detalla la arquitectura utilizada para el diseño del sistema predictivo y su conexión con la plataforma Moodle, lo cual se ilustra en la Figura 11. La interfaz del sistema consiste en la integración de elementos gráficos para interactuar con los usuarios, como botones, listas, cuadros de diálogo, cuadros de texto, entre otros. Los módulos de control y de lógica de la aplicación orquestan el funcionamiento del sistema para las funciones de predicción y seguimiento implementadas.

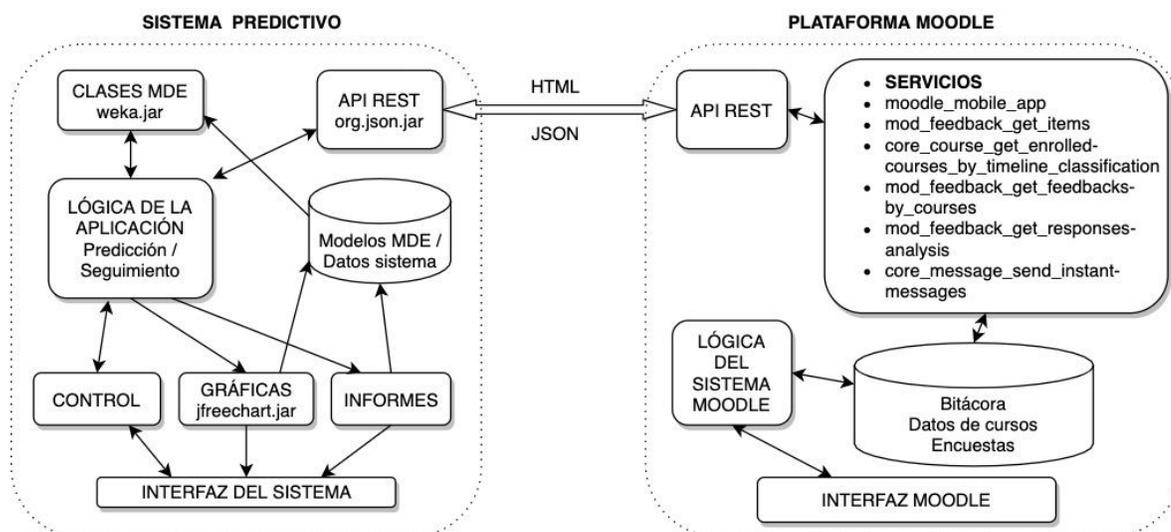


Figura 11 Arquitectura del sistema predictivo.

Para comunicarse y recuperar datos del sistema Moodle se utiliza su API REST, la cual ofrece un conjunto de servicios web para acceder a datos específicos de la plataforma, el formato de intercambio de datos utilizado fue JSON. Una vez recuperada la información, el módulo con las clases de MDE recupera los modelos de los algoritmos y realizar las predicciones de riesgo académico, las cuales son presentadas a los usuarios en formato de lista. Para el módulo de seguimiento, de igual forma, se accede a las encuestas en Moodle, se calculan los valores estadísticos y se generan las gráficas para mostrar los resultados. Los reportes y las imágenes son almacenadas en el directorio de trabajo de la aplicación.

### 6.3 Recuperación de Datos de Encuestas desde Moodle

Para acceder a los datos de las encuestas se necesita establecer una conexión al servidor que contiene el curso (ver Figura 12). Se deben contar con credenciales de un usuario con rol de profesor, para generar un token mediante el cual se autoriza el acceso a las funciones disponibles en la API de Moodle. Es decir, cada vez que se invoca una función, es necesario enviar el token de autorización.



Figura 12 Conexión inicial a la plataforma Moodle desde el sistema predictivo.

Una vez autorizado el acceso, se selecciona el curso, se obtienen las encuestas disponibles y se elige aquella que fue diseñada para registrar los datos de los alumnos con los atributos necesarios para realizar la predicción. Al momento de seleccionar la encuesta se recupera la información contenida en ella, como se muestra la Figura 13, en la que se observan algunas de las variables empleadas en la predicción de riesgo académico.

...	Carrera	ICNE	DIAG	Exani	Edad	Prepa	Resp
...	LCC	1200	50	770	23	UADY	MADRE
...	LCC	1222	62	841.5	18	UADY	MADRE
...	LCC	1050	40	584	18	No UADY	PADRE
...	LCC	1150	63	761.25	19	No UADY	MADRE
...	LCC	1216	70	864.5	18	UADY	MADRE
...	LCC	1210	46	767.5	21	No UADY	MADRE
...	LCC	1240	71	896.25	20	No UADY	MADRE
...	LCC	1198	69	829	20	No UADY	PADRE
...	LCC	1204	51	779.25	18	UADY	MADRE
...	LCC	1108	58	693.5	18	No UADY	PADRE

Figura 13 Selección de curso y encuesta Moodle para la recuperación de datos.

## 6.4 Selección de Modelos Predictivos

La interfaz del programa permite seleccionar entre tres modelos predictivos. Al elegir un modelo, por ejemplo Logistic (ver Figura 14), de forma predeterminada se asocian dos archivos que corresponden a los modelos generados para la clase Riesgo\_P y Riesgo\_A. Es posible seleccionar y asociar otros archivos de modelos generados con otros datos utilizando el botón seleccionar nuevos, se debe tener cuidado de que los archivos elegidos correspondan al tipo de modelo que se desea utilizar.



Figura 14 Selección de modelos predictivos a utilizar.

## 6.5 Predicción de Riesgo

Una vez recuperados los datos y elegido el método de predicción, será posible predecir el riesgo mediante el botón correspondiente. Internamente, se abren los archivos de los modelos a utilizar y los datos recuperados son procesados por los algoritmos de predicción para generar los resultados.

Desde el sistema se cargan las librerías requeridas y se preparan los datos para utilizar un objeto WEKA que cuenta con los métodos o algoritmos para la predicción. En el código fuente, se usa el método *read* de la clase *SerializationHelper*, para leer y cargar el archivo con el modelo seleccionado:

```
fmodelP="J48_2016_2019_Riesgo_P.model"
```

```
J48 treeP = (J48) SerializationHelper.read(new FileInputStream(fmodelP));
```

Se invoca al método que realiza la predicción de la instancia del alumno:

```
result = treeP.classifyInstance(instance2);
```

Si el resultado es igual a 0.0, significa que la clase “SI” es verdadera, en nuestros modelos significa la presencia de riesgo académico. En caso contrario, no hay riesgo académico.

El programa realiza de forma automatizada e iterativa la lectura de cada instancia recuperada y calcula su predicción. Los resultados de las predicciones, para cada alumno y con cada una de las variables de clase RIESGO\_A y RIESGO\_P, se muestran en la Figura 15.

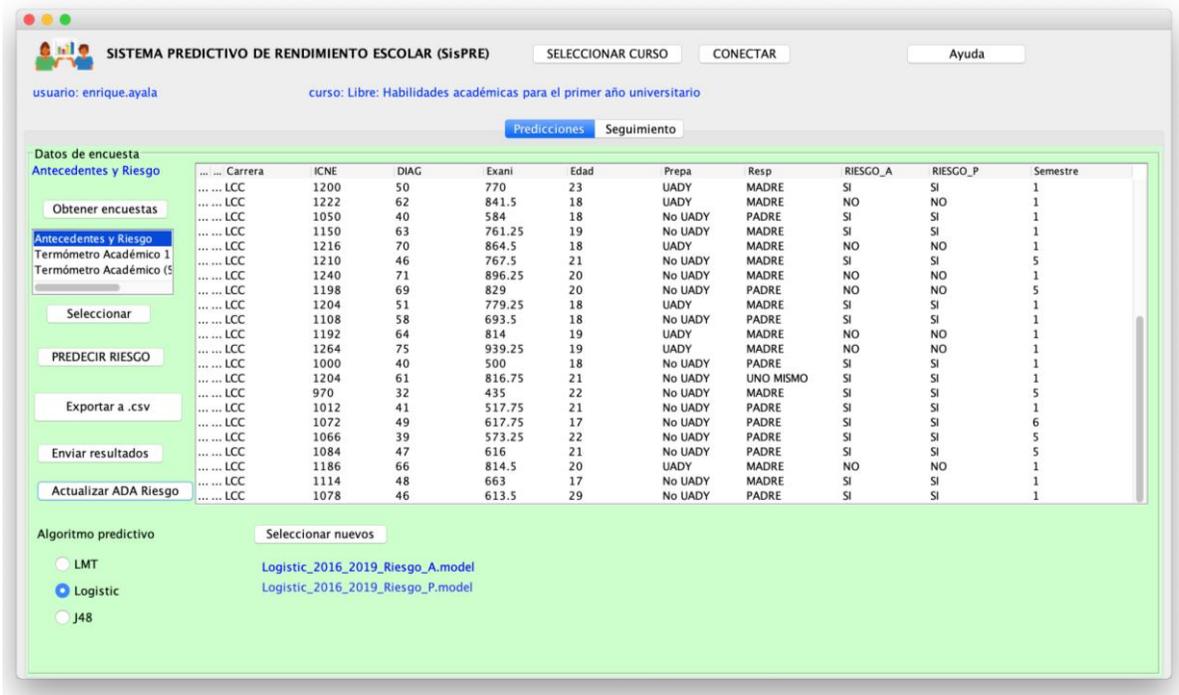


Figura 15 Resultado de la predicción mediante el algoritmo Logistic.

## 6.6 Comunicación de los Resultados

Una vez realizada la predicción, se cuenta con las siguientes opciones para utilizar o comunicar los resultados:

- Exporta a .csv: genera un archivo con formato separado por comas, con la lista de instancias, sus atributos y los resultados de la predicción.
- Enviar resultados: envía a cada alumno una alerta a través del sistema de mensajería de la plataforma Moodle, con el resultado de la predicción. Los avisos también se envían a los correos registrados de los alumnos. Un ejemplo de los mensajes enviados se muestra en la Figura 16.

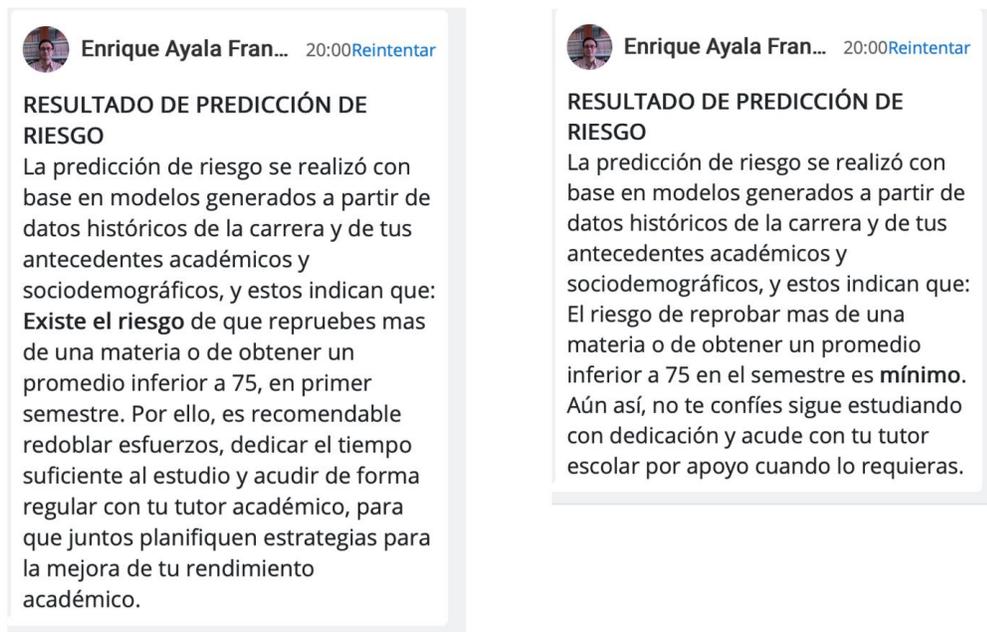


Figura 16 Mensajes con resultados de predicción enviados a los alumnos

- Actualizar ADA Riesgo: establece una comunicación con la plataforma Moodle para modificar la calificación o estado de alguna actividad tipo tarea diseñada previamente en el curso. Por ejemplo, un valor 0 podría indicar sin riesgo, mientras que un valor 10 indicaría riesgo académico; entonces, mediante las opciones de personalización para la presentación de recursos

en Moodle, se puede emplear este valor como requisito para mostrar o no información, restringiendo el acceso de manera condicional. En la Figura 17, se muestran opciones para restringir el acceso a una sección del curso, la cual muestra material adicional sólo a alumnos pronosticados con riesgo.

Restringir acceso

Restricciones de acceso

Estudiante debe de coincidir con lo siguiente

Calificación R\_A

debe ser  $\geq$  5 %

debe ser  $<$  %

Añadir restricción...

Figura 17 Restricción de acceso a recurso Moodle con base en calificación de una actividad.

## 6.7 Módulo de Seguimiento a Cursos

Para conocer la opinión de los alumnos respecto al desarrollo de sus clases y su rendimiento durante el semestre se diseñó un instrumento en la plataforma Moodle, con 5 preguntas (ver Figura 18), que fue aplicado de manera periódica en diferentes momentos del semestre.

1. La pertinencia del contenido, acorde a los objetivos de aprendizaje a sido: 1

(1)Muy Deficiente  (2)Deficiente  (3)Regular  (4)Bueno  (5)Excelente

2. El Método y estrategias de aprendizaje han sido: 1

(1)Muy Deficiente  (2)Deficiente  (3)Regular  (4)Bueno  (5)Excelente

3. Mi motivación en el desarrollo de actividades a sido: 1

(1)Muy Deficiente  (2)Deficiente  (3)Regular  (4)Bueno  (5)Excelente

4. Mi desempeño como alumno a sido: 1

(1)Muy Deficiente  (2)Deficiente  (3)Regular  (4)Bueno  (5)Excelente

5. La comunicación con el profesor(a) a sido: 1

(1)Muy Deficiente  (2)Deficiente  (3)Regular  (4)Bueno  (5)Excelente

Figura 18 Ítems de la encuesta de satisfacción para seguimiento periódico.

El módulo de seguimiento del sistema permite recuperar los datos de las encuestas almacenadas en el curso, calcular valores estadísticos de las respuestas aportadas por los alumnos y, finalmente, mostrar mediante gráficos el resumen de los resultados (ver Figura 19).

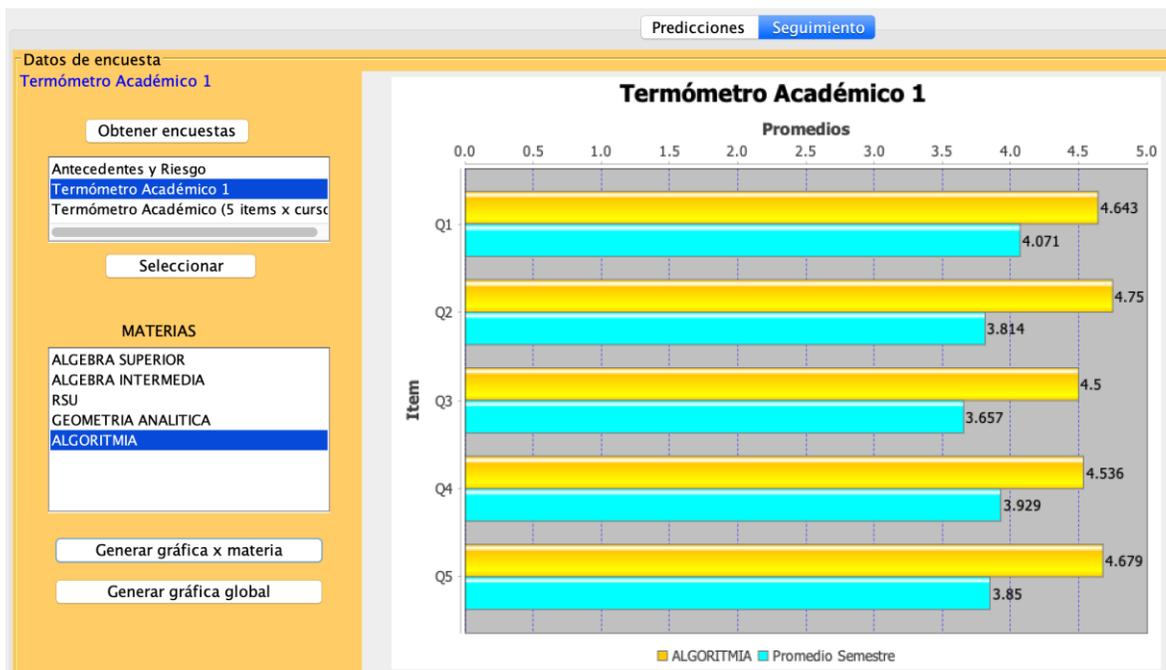


Figura 19 Módulo de seguimiento de satisfacción de los cursos.

Para generar las gráficas, se selecciona alguna de las materias de la lista y se presiona el botón de generar gráfica. La imagen correspondiente se mostrará y a la vez, se generará un archivo de la imagen en formato jpg. En la Figura 19 se muestran los resultados para una materia mediante un gráfico de barras, por cada pregunta la barra indica el promedio de las respuestas para la asignatura y se compara con el promedio de todas las respuestas de todas las materias.

El botón generar gráfica global, calcula el promedio general de todas las preguntas por cada materia, este valor resume el nivel de satisfacción de los alumnos por cada asignatura. En la Figura 20, se muestra un comparativo de los resultados globales

de todos los cursos, con los valores obtenidos se pueden determinar las materias con menor y mayor nivel de satisfacción.

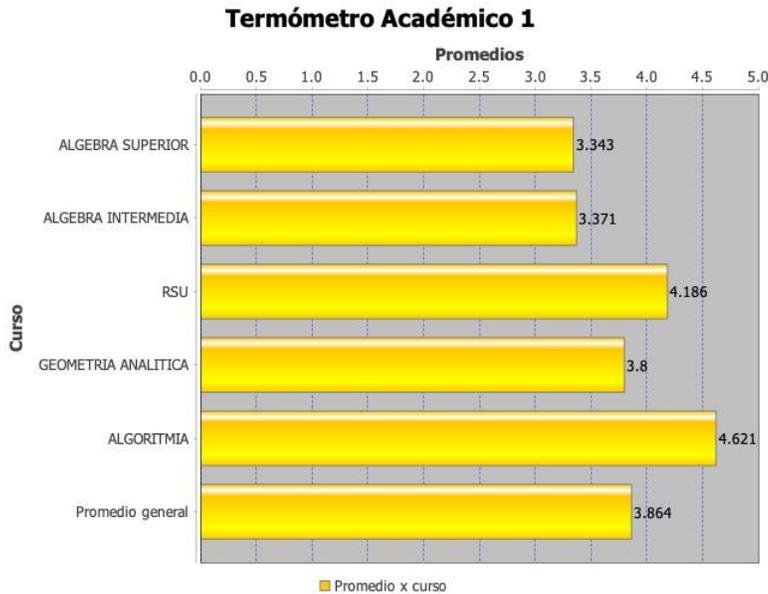


Figura 20 Resultados comparativos globales de los cursos.

Los resultados de este módulo ofrecen un punto de referencia para analizar el desarrollo de las asignaturas y el nivel de satisfacción de los alumnos. Las imágenes se utilizan para preparar informes los cuales serán compartidos con los profesores titulares de las asignaturas

## 6.8 Comentarios Finales

El módulo de predicción del sistema, además de pronosticar el rendimiento académico de estudiantes de nuevo ingreso permite la notificación personalizada de los resultados a los alumnos, lo cual en sí mismo ya es una estrategia de intervención puesto que brinda información para que los estudiantes en riesgo puedan tomar medidas preventivas respecto a sus estudios y su desempeño académico. También, la información es utilizada, por parte de profesores, tutores y personal directivo, para tomar medidas oportunas de apoyo e implementar otras estrategias de intervención educativa que se consideren viables, como el diseño de

curso o talleres para atender las necesidades de formación más urgentes de los alumnos.

El módulo de seguimiento ha permitido detectar situaciones de baja motivación y cursos con algunas problemáticas de desarrollo. Lo cual se ha comentado con los tutores y con la Secretaría Académica, y se han tomado algunas medidas para dar seguimiento y solución a las problemáticas detectadas, por ejemplo, pláticas de los alumnos con sus tutores e invitación para que participen en talleres de desarrollo de habilidades matemáticas o de autodirección del aprendizaje.

## 7. PROCESO DE INVESTIGACIÓN-ACCIÓN

El contenido de este capítulo presenta una descripción de las tareas realizadas durante el proceso de investigación-acción, mediante el cual se llevó a cabo la intervención educativa. La Figura 21 muestra las principales fases en las que quedaron enmarcadas las actividades que dieron forma al plan de acción para atender las necesidades detectadas. Inicialmente, se describe como se realizó el análisis diagnóstico, sus resultados y las necesidades de los alumnos que debían ser atendidas. A partir de los hallazgos, la etapa de planificación muestra el diseño del plan de intervención y las principales líneas de acción que se debían seguir. Posteriormente, se describe el proceso de implementación y seguimiento del plan, para concluir con un análisis cuantitativo y otro cualitativo de los resultados, que dan cuenta del grado de éxito de la intervención.



Figura 21 Pasos generales de la investigación-acción.

### 7.1 Análisis Diagnóstico

El propósito de esta etapa de la I-A, fue identificar y describir algunos de los principales factores que afectan el rendimiento académico de los estudiantes de nuevo ingreso de las carreras del área de cómputo en la Facultad de Matemáticas. Por otro lado, se busca conocer las condiciones actuales en las que se desenvuelven los estudiantes y las características principales de los programas de estudios, lo que permitirá identificar fortalezas y debilidades. El análisis de estos elementos es necesario para facilitar el diseño del plan de intervención y la definición de los mecanismos de apoyo para los estudiantes, por ejemplo,

impartición de talleres específicos en atención a requerimientos concretos, mejoras en los programas de tutoría y asesoría, pláticas, entre otros. Es decir, la información sobre los elementos que pudieran influir en el proceso educativo puede aportar una visión más amplia de las problemáticas de los estudiantes y de la escuela, lo que justificará las medidas adoptadas en el plan de mejora e intervención con un enfoque holístico y multifactorial.

Para realizar el análisis se contaba con información de las calificaciones de las asignaturas de primer semestre de alumnos inscritos entre 2016 y 2018. Sin embargo, consideramos necesario aplicar un cuestionario diagnóstico para recabar y conocer de forma directa la opinión de los alumnos respecto a varios elementos de su entorno académico y no académico, por ejemplo, sobre sus antecedentes, la infraestructura disponible, los servicios y apoyos que ofrece la escuela, las estrategias docentes, estrategias con apoyo de tecnología educativa, asesorías y tutorías, entre otros.

### ***7.1.1 Enfoque Metodológico para el Diagnóstico***

El estudio tiene un alcance exploratorio y descriptivo, está basado en datos cuantitativos y cualitativos (Hernández, Fernández y Baptista, 2014). Se definieron dos etapas para el desarrollo de las actividades. La primera, para analizar los resultados académicos históricos de los estudiantes de las carreras de computación, con lo cual se podrá dimensionar la magnitud de la problemática y posiblemente identificar grupos específicos con mayor urgencia de intervención. La segunda etapa, con el propósito de identificar probables causas de bajo rendimiento académico y perfilar posibles estrategias de intervención.

Participantes. Para el primer análisis se consideraron datos históricos de 392 alumnos de primer ingreso (2016-2018), distribuidos como sigue: 77 alumnos de Licenciatura en Ciencias de la Computación (LCC), 90 alumnos de Licenciatura en Ingeniería en Computación (LIC) y 224 alumnos de Licenciatura en Ingeniería de Software (LIS). En el segundo momento, para la aplicación de la encuesta

diagnóstica (ver Anexo 1), se contó con la participación de 116 alumnos de nuevo ingreso inscritos en el semestre agosto-diciembre de 2019. De estos, la cantidad de alumnos por carrera fue variada: 29 de LCC, 35 de LIC y 52 de LIS.

Procedimiento. En la primera etapa se recuperó, de los registros de control escolar institucional, la información académica disponible de la población de interés, específicamente los datos correspondientes a las calificaciones de 5 asignaturas de primer semestre, posteriormente se procedió a su depuración, sistematización y análisis estadístico generando los informes y gráficas correspondientes. En la segunda etapa, se diseñó un instrumento diagnóstico para identificar factores y posibles estrategias para la mejora del rendimiento académico. El cuestionario fue diseñado en la plataforma *Google Forms* para su aplicación en línea y para facilitar la consulta de resultados en forma tabulada y gráfica, y para su exportación a diferentes formatos de archivo. Se aplicó 15 días previos a la conclusión del período escolar, lo anterior con la debida autorización de la dirección de la escuela y el consentimiento de profesores y alumnos, asegurando en todo momento la confidencialidad de la información. El cuestionario contiene 29 ítems, la mayoría para responder en escala Likert de 5 puntos, y otras preguntas fueron de tipo ensayo y de selección. Los datos recabados se procesaron y analizaron mediante el software de hojas de cálculo (Excel) y con el programa factor. Fueron generados reportes de estadística descriptiva y se verificó la validez interna de la encuesta. En el caso de los datos cualitativos, se recuperó y organizó la información, luego se realizó un análisis de contenidos para identificar las categorías de interés.

### ***7.1.2 Resultados del Diagnóstico***

**Resultados de los datos históricos.** Los resultados de los datos los alumnos de primer semestre que ingresaron entre 2016 y 2018. La información fue obtenida de los registros de control escolar y datos proporcionados por secretaría académica, de alumnos pertenecientes a las carreras del área de computación. Se contó con un total de 392 registros. El promedio general de calificaciones es de 68.5.

Los resultados generales se presentan en la Figura 22. La primera gráfica muestra que un 44% de los alumnos obtuvo un promedio de calificación entre 0 y 69, lo que corresponde a un resultado de No acreditado en el sistema institucional. También, en la segunda gráfica sobre materias reprobadas, se puede ver que el 61% de los alumnos reprueban al menos una materia y el 44% reprueba dos o más.

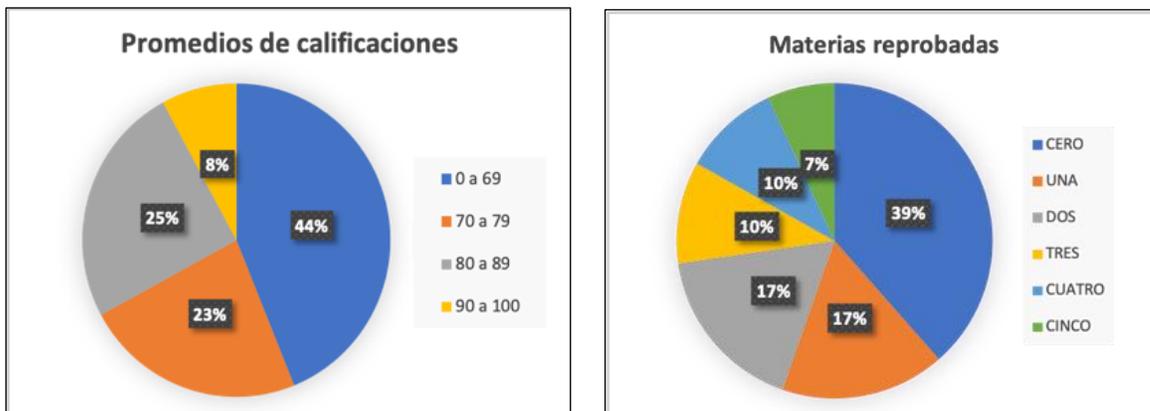


Figura 22 Resultados generales de promedios de calificaciones y materias reprobadas de las carreras de computación (2016-2018).

La Figura 23 muestra los promedios de calificaciones y materias reprobadas correspondientes a la carrera LIC.

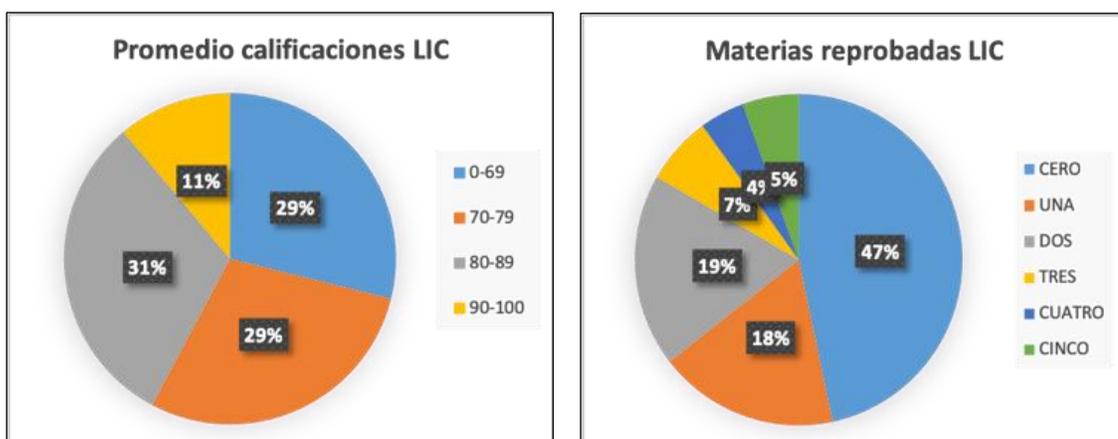


Figura 23 Resultados de promedios de calificaciones y materias reprobadas LIC (2016-2018).

En la Figura 24 se observan los valores de promedios de calificaciones y materias reprobadas de la carrera LIS.

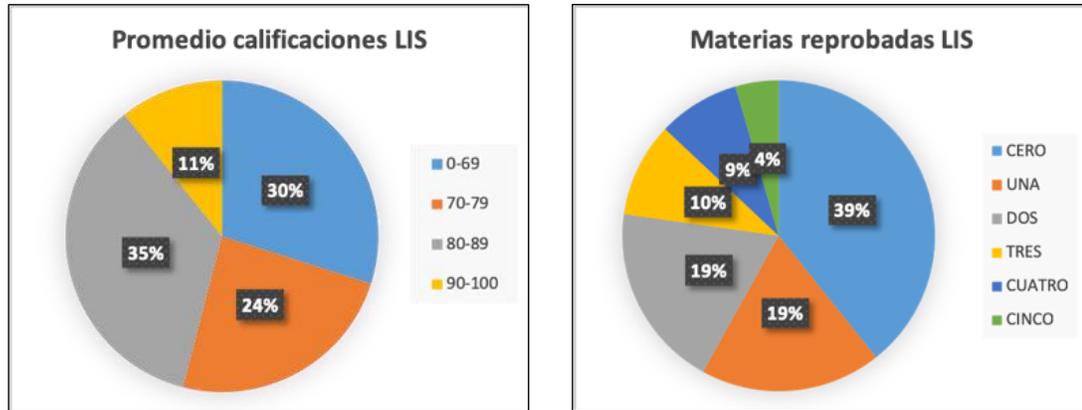


Figura 24 Resultados de promedios de calificaciones y materias reprobadas LIS (2016-2018).

En la revisión de resultados por carreras, para los programas de LIC y LIS como se muestra en la Figura 23 y Figura 24, los resultados son muy similares, ya que aproximadamente un 30% de alumnos, en ambos casos, obtiene un promedio de calificación No aprobatoria. En cuanto a materias reprobadas, un 53% de LIC y un 61% de LIS reprueban una o más asignaturas.

La Figura 25 muestra los resultados para la carrera LCC, en el período 2016 – 2018.

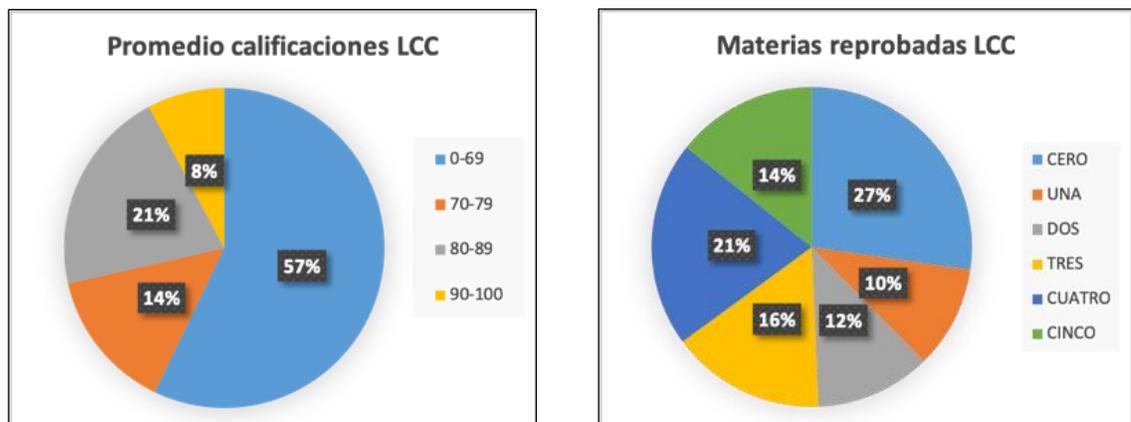


Figura 25 Resultados de promedios de calificaciones y materias reprobadas LCC (2016-2018).

En el caso de LCC, Figura 25 primera gráfica, se muestra que un 57% de los alumnos tiene un promedio de calificaciones No aprobatorio. En la segunda gráfica, sobre reprobación, se presentan porcentajes de materias reprobadas más altos comparado con los otros programas, ya que un 74% de los alumnos reprueba una o más asignaturas, o el 64% reprueba dos o más. Se observa que esta carrera presenta altos índices de reprobación con 3, 4 y hasta 5 asignaturas.

**Resultados de la encuesta.** Para la interpretación de la encuesta en las preguntas de escala Likert, las categorías fueron: 1-Totalmente de acuerdo, 2-De acuerdo, 3-Indeciso/no sé, 4-En desacuerdo, 5-Totalmente en desacuerdo. Se obtuvo un Alpha de Cronbach de 0.72, que se encuentra dentro del límite de 0.7 a 0.9 que indica una consistencia interna aceptable del instrumento para la escala Likert de 5 puntos.

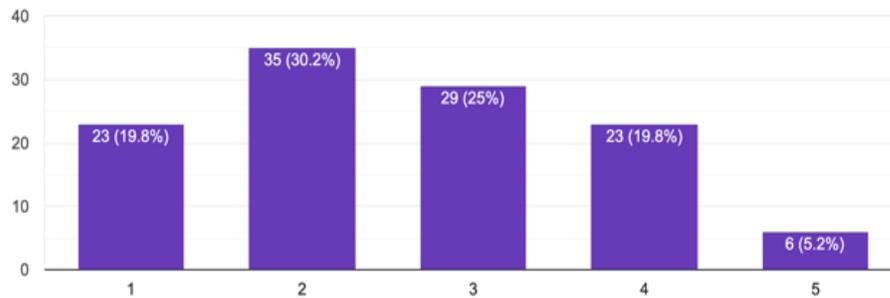
Se presentan los resultados más relevantes derivados de la aplicación de la encuesta diagnóstica, en particular aquellos aspectos que pudieran requerir acciones de intervención o mejora. Respecto a las expectativas de promedio de calificaciones al finalizar el semestre, se tiene que un 32.8% de los alumnos espera obtener un promedio entre 70 y 79, y sólo un 1.7% un promedio inferior a 70. Esto difiere con los datos históricos, en donde el 23% de los alumnos obtuvo un promedio entre 70 y 79 y un 44% tiene un promedio inferior a 70.

En cuestión de expectativas de materias reprobadas el 40% de los estudiantes piensa que reprobará sólo una materia y un 32.8% considera que reprobará dos o más. Históricamente, el 17% de los alumnos reprueban sólo una materia, pero el 44% reprueba 2 o más.

Referente a las condiciones del estudiante al ingresar, la mayoría considera que se sentían preparados para realizar sus estudios (86%), mientras que sólo un 14% opinó que no estaba bien preparados.

La Figura 26 muestra aspectos relacionados con antecedentes y expectativas de los estudiantes. Sobre el nivel de conocimiento y exigencias académicas (ítem 5), un 50% piensa que es más alto de lo que esperaban.

5. Siento que el nivel de conocimientos y exigencias de los cursos es más alto del que yo esperaba.  
116 respuestas



6. Considero que tengo buenos hábitos de estudio, lo que me permite estudiar de manera autónoma y responsable.  
116 respuestas

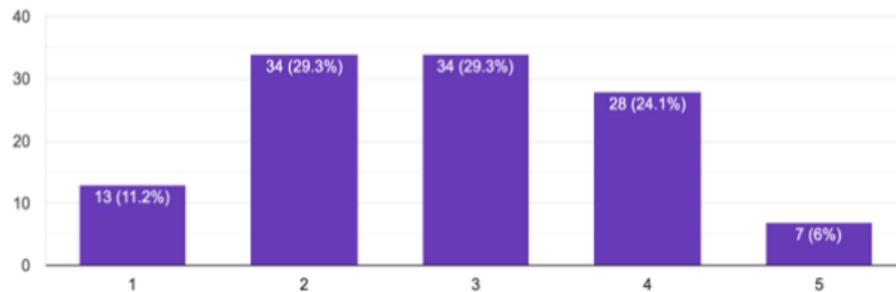


Figura 26 Opinión de los estudiantes sobre antecedentes y expectativas personales.

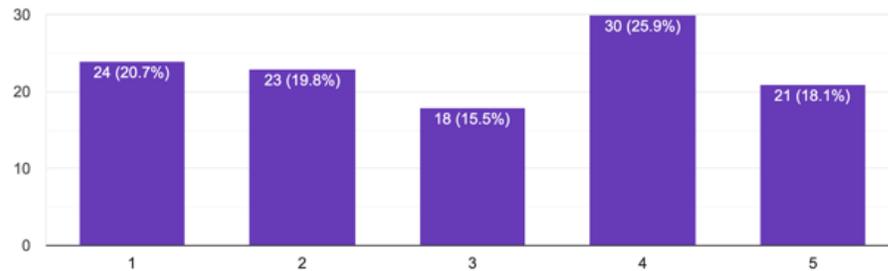
En el tema de hábitos de estudio (ítem 6), un 30.1% considera que no tiene buenos hábitos de estudio, y un 29.3% que ni buenos ni malos.

La Figura 27 presenta las opiniones que reflejan las situaciones personales de los estudiantes, aspectos que les están afectando o la necesidad de apoyos por parte de la institución educativa.

En situaciones personales (ítem 7), un 40.5% que considera que hay situaciones que le afectan en este momento en su rendimiento y un 31.9% considera que necesita más apoyo por parte de la escuela (ítem 8).

7. Creo que estoy pasando por situaciones personales, familiares o económicas que están afectando mi rendimiento académico.

116 responses



8. Pienso que necesito más apoyo por parte de la escuela para mejorar mi desempeño académico.

116 responses

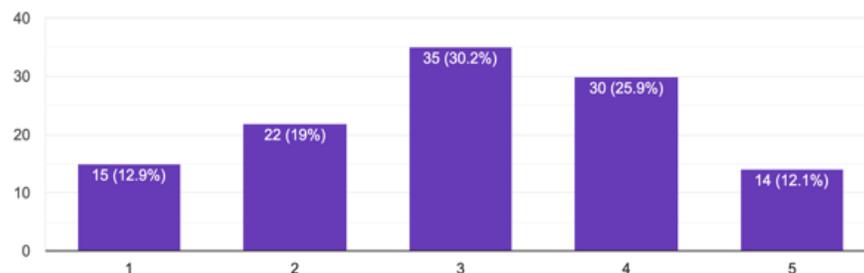


Figura 27 Opinión de los estudiantes sobre situaciones personales y apoyos requeridos.

En cuanto a preparación de profesores, materiales de enseñanza, métodos de enseñanza, estrategias de evaluación, evaluación formativa; la mayoría de las opiniones son: de acuerdo o completamente de acuerdo. Es decir, en cuanto a estrategias docentes se tiene una opinión positiva.

Las opiniones sobre empleo de la plataforma virtual, variedad de herramientas usadas en la plataforma y uso de tecnología educativa básica, tienden hacia valores satisfactorios. Sólo en uso de tecnología educativa avanzada (ver Figura 28), las opiniones son en el sentido de poco empleo de esta.

18. Los profesores utilizan en sus clases tecnología educativa novedosa o avanzada (clases invertidas, aprendizaje móvil, videos interactivos,....uegos educativos, herramientas colaborativas, etc.)  
116 responses

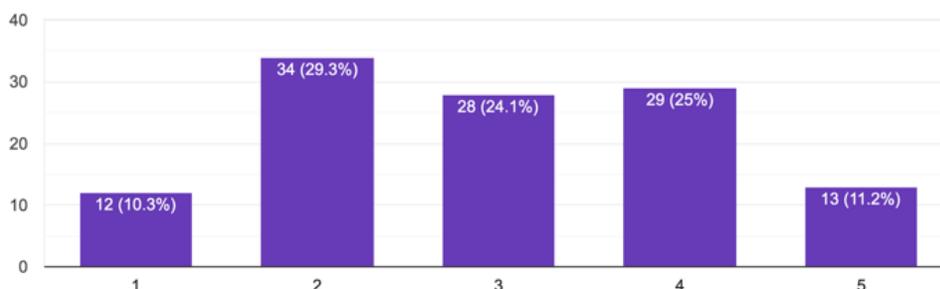


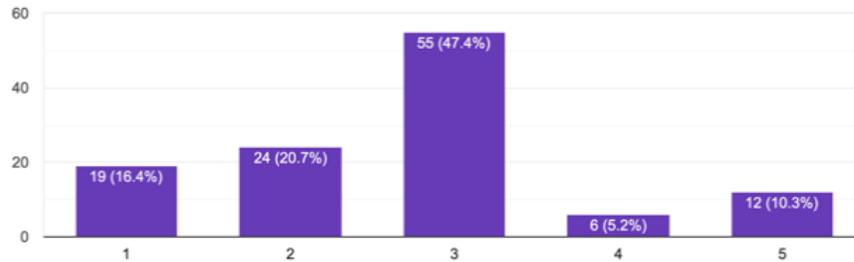
Figura 28 Opinión de los estudiantes sobre uso de tecnología educativa novedosa.

En el tema del entorno institucional, en general, hay impresiones positivas en cuanto a los recursos físicos e infraestructura disponible y consideran que los planes de estudios están diseñados adecuadamente. Sobre el acceso a becas y el desarrollo de talleres de apoyo, en la mayoría de los casos se está de acuerdo en que si ayudarán para mejorar el rendimiento académico.

En la Figura 29 se observa que la tutoría tiene una tendencia muy marcada por la respuesta 3 (indeciso/no sé), lo que indica que no tienen claro como la tutoría les puede ayudar de forma determinante en sus estudios. Se complementa esta información con el ítem 26, en dónde se pone de manifiesto que la mayoría de los alumnos no asisten o asisten solo a una o dos sesiones de tutoría al semestre, por lo cual no se tiene una influencia significativa en el rendimiento escolar de los estudiantes. Prácticamente la misma tendencia se presentó para el caso de la asesoría académica.

25. Las sesiones de tutoría (por dudas administrativas o académicas, apoyo carga de materias, procesos administrativos, etc., con un tutor así...an ayudado para mejorar mi desempeño académico.

116 respuestas



26. Las veces que he acudido a sesiones de tutoría con mi tutor(a) asignado son:

116 respuestas

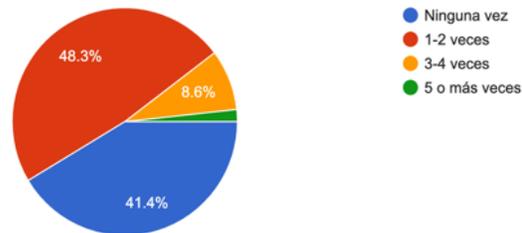


Figura 29 Opinión sobre el programa de tutorías institucional.

**Opiniones para la Mejora.** Los alumnos hicieron comentarios para mejorar varios aspectos de la institución, así como algunos tipos de apoyos que la escuela podría brindar a los alumnos para mejorar su desempeño (Tabla 8 y Tabla 9).

Tabla 8 Principales apoyos requeridos por los estudiantes.

APOYOS REQUERIDOS
Becas, alimentos, apoyo a alumnos foráneos.
Asesorías, apoyo académico, mejores horarios de asesorías, más sesiones.
Tutoría, apoyo psicológico o psiquiátrico, flexibilidad de horario, facilidad para elegir tutor, más sesiones.
Mejorar pedagogía, retroalimentación oportuna, mejorar planeación didáctica, material de apoyo (manuales, prácticas), diferentes formas de evaluar.
Talleres extracurriculares, clases de apoyo en matemáticas.
Saber cómo organizarme y responsabilidad personal, dedicación y disciplina, mejorar habilidades de aprendizaje, manejo de prioridades.

Además de los apoyos económicos, también se hizo notar la necesidad de mejorar los apoyos académicos que se reciben en el primer tramo de sus carreras. Las tutorías y asesorías fueron mencionadas de forma frecuente, así como el desarrollo de talleres específicos.

Tabla 9 Aspectos a mejorar en el entorno de la escuela.

<b>ASPECTOS A MEJORAR</b>
Mejorar pedagogía, mejorar planeación didáctica, retroalimentación oportuna, materiales de apoyo (manuales, prácticas), diferentes formas de evaluar, valor y duración de los exámenes, tiempos asignados a actividades.
Actualizar temas en varias asignaturas, tendencias en tecnología y lenguajes de programación.
Enfoque práctico de materias en primer semestre, relacionado con la profesión.
Más y mejores computadoras en centro de cómputo, disponibilidad de software, mejores muebles y salones.
Aires acondicionados, lugares para estudiar y descansar, espacios al aire libre, instalaciones deportivas.
Materiales para laboratorios, proyectores con HDMI.
Libros en biblioteca, más cubículos para estudio, limpieza de baños e instalaciones.

En cuanto a los aspectos a mejorar, se destacan cuestiones pedagógicas que tienen que ver con la forma como actualmente se imparten las clases por algunos maestros, la actualización de contenidos de las materias y temas relacionados con infraestructura y recursos.

### **7.1.3 Discusión de los Resultados del Diagnóstico**

Considerando que los participantes de la encuesta son alumnos de primer ingreso y aún están en un proceso de transición y adaptación a la universidad, se observa que sus expectativas son optimistas respecto a los logros que consideran alcanzar. Sin embargo, los resultados generales derivados del análisis de datos históricos

muestran que aproximadamente un 44% de los alumnos tienen problemas de rendimiento académicos y reprobación (considerando que reprueban 2 o más asignaturas), por lo que ese conjunto de alumnos enfrentará problemáticas referentes a su tránsito escolar y los coloca en una situación de riesgo académico, lo que puede derivar en deserción o en baja de la escuela por cuestiones reglamentarias, si no logran aprobar sus asignaturas. Particularmente, la carrera de LCC tiene un porcentaje de materias reprobadas más alto que los otros programas, ya que el 64% de los alumnos reprueba dos o más asignaturas.

En general, las respuestas reflejaron mayormente opiniones positivas respecto a la disponibilidad de recursos físicos, el diseño de los planes de estudio, la preparación de los maestros y las estrategias docentes que emplean, aunque se observa poco uso de tecnología educativa novedosa o avanzada. Sin embargo, en las preguntas abiertas se recibieron varias sugerencias para la mejora de varios de estos elementos.

Un porcentaje importante de alumnos siente que el nivel de exigencia es más alto del que imaginaban y sus hábitos de estudio no son del todo adecuados para lidiar con estas exigencias, además de que existen situaciones de índole extraescolar que les afectan, por lo que necesitan más apoyo por parte de la Institución. El apoyo que requieren es de tipo económico, pero también académico a través de asesorías y tutorías, o mediante talleres o cursos que les permitan regularizarse en temas específicos o para desarrollar sus habilidades de organización y aprendizaje; y se mejore su nivel de responsabilidad y disciplina.

Hay aspectos de los planes y programas de estudio que se tienen que revisar para el buen desarrollo de las asignaturas y el despliegue de estrategias pedagógicas efectivas. En particular, se menciona el tema de actualizar los contenidos de las asignaturas de primer semestre con elementos recientes y relacionados con su profesión, además de diseñarlas con un enfoque más práctico, lo cual posiblemente

podiera ser un factor que incremente la motivación de los grupos y los enganche para que se comprometan con el estudio.

Las asesorías y tutorías se consideran un apoyo importante para los alumnos, sin embargo, la forma de operar el programa de tutorías y asesorías en la actualidad hace que la asistencia a estas actividades sea muy baja, por lo que no se tiene una influencia determinante para incidir en los resultados escolares. Esto se relaciona con la problemática de desconocimiento de antecedentes e intervención oportuna, pues además de no contar con información para diagnosticar posibles casos de alumnos en riesgo, tampoco se cuenta los mecanismos que pudieran coadyuvar a resolver la problemática del bajo rendimiento académico.

Respecto a factores que inciden en el rendimiento, se han detectado algunas áreas de mejora en aspectos pedagógicos, de la operación de los planes de estudio y de los servicios, por lo cual se podrían empezar a definir estrategias específicas para atenderlos. Se requiere atender las situaciones personales y mejorar los hábitos de estudio, especialmente es necesario apoyarlos para que logren una pronta transición y adaptación a la universidad, lo cual se podría realizar mediante la mejora del sistema de asesorías y tutorías, la impartición de talleres o cursos que atiendan las necesidades específicas de los alumnos.

#### ***7.1.4 Conclusiones del Análisis Diagnóstico***

A partir de este estudio se detectaron con mayor precisión grupos de alumnos con bajo rendimiento académico en los primeros semestres, y elementos de la institución susceptibles de ser mejorados. Con la información disponible se diseñarán estrategias de intervención educativa para atender las necesidades más apremiantes de los alumnos.

Los resultados derivados de este estudio permitieron caracterizar a los alumnos de las carreras del área de computación, y conocer varios elementos del entorno educativo que influyen en su desarrollo académico. Al cruzar la información de los

datos históricos con los resultados obtenidos en la encuesta diagnóstica, se abordó la problemática desde diversas perspectivas y se logró un análisis más amplio y profundo con lo cual se podrán planificar estrategias de intervención centradas en las necesidades más relevantes y urgentes de los alumnos. Dichas estrategias de intervención educativa tendrán el propósito de modificar o mejorar el rendimiento académico de los estudiantes en riesgo de fracaso escolar.

## **7.2 Planeación de la Intervención Educativa**

Para atender las problemáticas educativas existen diversas modalidades de intervención (Zeng et al., 2018). También se ha visto que intervenciones oportunas pueden mejorar procesos cognitivos, como la autoeficacia, la autorregulación, la autodeterminación, favorecer la metacognición y las habilidades que se requieren para la vida académica y el desarrollo profesional. Adicionalmente, se ha encontrado que el uso de las TIC puede impactar positivamente en la efectividad y alcance del aprendizaje, además de eliminar barreras de espacio y tiempo de instrucción, así como brindar facilidades para la comunicación, una mayor flexibilidad y motivación (Benítez, Barajas y Hernández, 2014).

Los resultados generales encontrados por Masino y Niño (2016), en la revisión de varias intervenciones, indican que las intervenciones para mejorar el rendimiento y el aprendizaje de los estudiantes, son más efectivas cuando se emplean dos o más elementos que generan algún tipo de cambio en los procesos educativos. Tomando en cuenta los aspectos mencionados previamente, y de manera central, basados en las necesidades detectadas de los estudiantes, se plantearon varias estrategias de intervención, con énfasis en la utilización de tecnología educativa como elemento facilitador de su implementación. Considerando los elementos de un plan de intervención descritos en Hernández et al. (2014) y Barraza (2010), se definieron los principales puntos del plan de intervención educativa a implementar.

### **7.2.1 Prioridades**

Se ha identificado una problemática de bajo rendimiento académico y reprobación, a partir del análisis de resultados en las calificaciones de alumnos de primer semestre de los programas educativos del área de computación. En el caso del programa Licenciatura en Ciencias de la Computación (LCC), un 57% de los alumnos tiene un promedio de calificaciones No Aprobatorio y presentan porcentajes de materias reprobadas más altos comparado con los otros programas ya que un 74% de los alumnos LCC reprueba una o más asignaturas, o el 64% reprueba dos o más.

Derivado del análisis diagnóstico, previamente descrito, y del uso de modelos predictivos generados con herramientas de la minería de datos educativas (ver el apartado proceso de la minería de datos educativos), fueron reconocidas otras problemáticas del contexto institucional y de los antecedentes de los alumnos. Los resultados mostraron que hay baja asistencia y deficiencias en el programa de tutorías y asesorías; también, hay la necesidad de apoyar a los estudiantes en su transición a la universidad, además de dar un seguimiento más puntual a su desempeño y adaptación a su nuevo entorno; prioritariamente, es necesario un mayor desarrollo y maduración de habilidades cognitivas para mejorar el aprendizaje de los alumnos de nuevo ingreso y el fortalecimiento de habilidades de autodirección.

### **7.2.2 Beneficios Esperados de la Intervención**

El desarrollo del proyecto de intervención atenderá las necesidades detectadas relativas a la transición de los estudiantes de bachillerato a estudios universitarios y el bajo rendimiento escolar. En el rendimiento académico intervienen múltiples factores personales, sociales, institucionales y de otra índole, que influyen en el proceso de aprendizaje y sus resultados (Montes y Lerner, citados en Monsalve, 2016). Las características y estructura del sistema escolar tienen un papel determinante para canalizar o contener la influencia de un entorno adverso o de

situaciones de vulnerabilidad en las que se pudieran encontrarse los estudiantes. Además, la universidad, a través de sus programas de estudio, debe ofrecer la infraestructura, apoyos y acciones necesarias para el desarrollo exitoso del proceso educativo y la formación integral de los estudiantes.

En el plano teórico y pedagógico, se analizará la influencia de nuevos diseños instruccionales basados en el uso de recursos TIC para el desarrollo de capacidades de autorregulación, motivación y metacognición del aprendizaje. Por ello, la propuesta de intervención tendrá el propósito de generar cambios actitudinales y cognitivos para que los alumnos puedan desarrollar sus habilidades académicas y afrontar de una mejor manera los retos del aprendizaje en sus primeros años de estudio universitarios.

### **7.2.3 Metas de la Intervención**

- Utilizar la información, modelos y sistemas, para la identificación temprana de alumnos en riesgo académico.
- Mantener un seguimiento del proceso educativo, para la mejora del ejercicio docente, para la planificación educativa y la toma de decisiones oportuna de docentes, directivos y alumnos.
- Comunicar resultados de detección de riesgo a alumnos e involucrados en el proceso educativo.
- Comunicar resultados de seguimiento a cursos en el semestre a los involucrados del proceso educativo.
- Implementar un curso con diseño pedagógico mediado por TIC para facilitar el aprendizaje y el desarrollo de habilidades académicas de los estudiantes.
- Elaborar recursos y materiales didácticos digitales para utilizarse durante la impartición de curso.

- Fomentar el desarrollo de habilidades cognitivas y académicas esenciales para los alumnos de nuevo ingreso que faciliten su adaptación y transición a la universidad.
- Evaluar el impacto de las medidas de intervención adoptadas.

#### **7.2.4 Situación Inicial**

El eje sobre el que se integraron la mayoría de las actividades de la intervención educativa fue el diseño e impartición de un curso donde se combinaron varios elementos tecnológicos y pedagógicos. El período de impartición fue del 1º de octubre de 2020 al 30 de enero de 2021.

Destinatarios. La intervención educativa está dirigida al grupo de estudiantes de LCC de nuevo ingreso. La anterior decisión, es debida a que en los últimos años han obtenidos los promedios de calificaciones más bajos y sus índices de reprobación son los más altos de entre todas las carreras en la Facultad de Matemáticas, donde se ubica este programa educativo. Los datos de los modelos predictivos también confirman tal situación de riesgo académico. Para el periodo escolar, en el que se realizó la intervención, se inscribieron 28 estudiantes a la carrera, 4 mujeres y 24 hombres, con edades entre 17 y 29 años, y media de 19.48 años.

Personas. El personal involucrado fue una profesora y un profesor (el autor de la investigación) para la impartición conjunta del curso en modalidad no presencial (e-learning). También participaron en otras acciones la responsable de la Secretaría Académica y el responsable del programa de tutorías de la Facultad, quienes estuvieron al tanto de la definición del diseño implementado y su seguimiento, además de que aportaron sugerencias para mejorar la propuesta. Así mismo, se contó con la participación de los tutores asignados de los alumnos.

Recursos para ejecutar el plan. Respecto a los recursos disponibles en la institución se cuenta con los siguientes: infraestructura de redes de computadoras con acceso

a Internet que cubre las necesidades de toda la Facultad de Matemáticas; se tienen servidores para instalar sistemas de gestión de cursos en línea, bases de datos o software de comunicaciones; la plataforma para la gestión de cursos virtuales Moodle está disponible a nivel local (en los servidores de la escuela) o a nivel institucional (en servidores en la nube de Microsoft); también, se tiene un convenio con Microsoft, por lo que están disponibles las aplicaciones de Office y la plataforma Teams. Adicionalmente, se utilizó el sistema predictivo y de seguimiento de rendimiento escolar desarrollado en la primera etapa del proyecto, para identificar a los alumnos en riesgo académico y para recabar opiniones sobre el desarrollo de los cursos.

Para la impartición de clases de manera presencial, se cuenta con 30 salones y 10 salas de cómputo con capacidad para atender a 300 alumnos de forma simultánea, el software requerido se encuentra instalado en los equipos. Todos los salones tienen cobertura para acceder a la red mediante WIFI o mediante instalaciones cableadas, además, la mayoría tiene instalado un proyector de forma permanente. Para la realización de actividades prácticas se cuenta con 3 laboratorios: de redes, de robótica y de electrónica.

La biblioteca del campus tiene acervo físico de cerca de 40,000 ejemplares, y ofrece cubículos para grupos de estudio, lugar para exposiciones, hemeroteca, videoteca, etc. También tiene recursos digitales, ya que pertenece a la CONRICYT y las bases de datos se pueden consultar de manera remota.

Dadas las condiciones ocasionadas por la pandemia, se realizó una encuesta por parte de la escuela, en la que se determinó que la mayoría de los maestros y los alumnos cuentan con los recursos tecnológicos (computadora o equipo con acceso a Internet, software de comunicaciones, etc.) que les permite acceder a los cursos en línea y a los recursos digitales disponibles en la institución.

### 7.2.5 Líneas de Acción

En la Tabla 10 se presentan las líneas de acción estratégicas y actividades, que fueron consideradas para el logro de las metas del proyecto de intervención. En función de los recursos materiales y humanos disponibles, el alcance de la intervención y el tiempo asignado para la implementación del plan y la modalidad en la que se diseñó el curso.

Tabla 10 Líneas de acción para la intervención educativa.

LÍNEA DE ACCIÓN	META	PROYECTO / ACTIVIDAD
1. Detección de alumnos en riesgo	1 y 3	<ul style="list-style-type: none"><li>- Diseño de cuestionario con datos de antecedentes de alumnos en la plataforma Moodle.</li><li>- Uso de sistema predictivo de rendimiento escolar, módulo de predicción.</li><li>- Notificación de resultados de riesgo mediante el sistema y por correo electrónico.</li><li>- Actualización de actividades en sistema Moodle para presentación condicional de materiales.</li></ul>
2. Seguimiento de cursos	2 y 4	<ul style="list-style-type: none"><li>- Diseño de encuestas de opinión en la plataforma Moodle.</li><li>- Uso del sistema para el seguimiento de cursos mediante encuestas de opinión, módulo de seguimiento.</li><li>- Notificación de resultados por correo electrónico.</li><li>- Sesiones con alumnos para platicar sobre sus avances o dificultades.</li></ul>
3. Diseño e impartición de curso para el desarrollo de habilidades académicas	5, 6 y 7	<ul style="list-style-type: none"><li>- Diseño instruccional y definición de la planeación didáctica.</li><li>- Elaboración de recursos digitales con contenido multimedia e interactivos.</li><li>- Diseño de actividades de aprendizaje mediadas por tecnología educativa, para el logro de los objetivos de aprendizaje.</li></ul>

		<ul style="list-style-type: none"> <li>- Implementación de los contenidos en la plataforma virtual.</li> <li>- Impartición de sesiones de clase síncronas con herramientas de videoconferencia.</li> <li>- Evaluación de alumnos y objetivos del curso.</li> </ul>
4. Medición de resultados de la intervención	8	<ul style="list-style-type: none"> <li>- Diseño de instrumentos para medir efectos de la intervención.</li> <li>- Recuperación y análisis de datos.</li> <li>- Presentación de informes y resultados.</li> </ul>

### **7.3 Implementación del Plan**

Las actividades planificadas se desarrollaron de septiembre de 2020 a febrero de 2021. En general, se pudieron cumplir todas las tareas programadas, el único contratiempo fue la suspensión de clases durante dos semanas debido a fenómenos meteorológicos que afectaron el suministro eléctrico en la región.

#### ***7.3.1 Detección de Alumnos en Riesgo***

Para registrar los datos con los antecedentes de los alumnos se diseñó un cuestionario en la plataforma Moodle solicitando la siguiente información: carrera, resultados en el examen de ingreso (ICNE, DIAG y Resultado total), edad, preparatoria de procedencia, el responsable del alumno y semestre. Los alumnos contestaron el instrumento la primera semana de clases.

El programa con el sistema predictivo se instaló y se ejecutó en la computadora del profesor. Una vez aplicada la encuesta, se utilizó el módulo de predicción para establecer conexión al curso en Moodle y recuperar los antecedentes de los alumnos. Con los datos recabados, se aplicaron los modelos y algoritmos predictivos de riesgo académico y se generaron informes con los resultados obtenidos.

La predicción de riesgo fue comunicada a los estudiantes enviando mensajes personalizados, utilizando el sistema predictivo y la plataforma de notificaciones de Moodle, que también reenvía los mensajes al correo electrónico de las cuentas asociadas de los estudiantes en la plataforma.

Adicionalmente, mediante el módulo de predicción, se actualizaron los valores o calificaciones de una actividad tipo tarea generada previamente en el curso en Moodle, la cual sirvió como criterio para mostrar de forma adaptativa o condicional algunos contenidos y apoyos adicionales para los alumnos, pronosticados con bajo rendimiento escolar por el sistema.

Se enviaron cartas a los tutores de los estudiantes pronosticados en riesgo académico, explicando el contexto y haciendo la invitación para que agendarán citas de apoyo y se les diera seguimiento a estos casos. De igual forma se informaron los resultados a la Secretaría Académica y se pidió espacio en un taller de desarrollo de habilidades matemáticas para inscribir a aquellos alumnos en riesgo que lo quisieran. Los tutores les recomendaron el tomarlo.

### **7.3.2 Seguimiento de Cursos**

La estrategia seguida para dar seguimiento a los avances y grado de satisfacción de los alumnos con las materias que estaban cursando en el semestre, fue la aplicación periódica de una encuesta de satisfacción con 5 preguntas para cada asignatura cursada, relativas a: 1) la pertinencia de los contenidos, 2) el método o estrategias de aprendizaje, 3) la motivación en el desarrollo de las actividades, 4) su desempeño como alumno, y 5) la comunicación con el profesor. La escala utilizada para expresar su satisfacción es de 5 niveles, desde Muy deficiente hasta Excelente. Los alumnos contestaron el instrumento en tres ocasiones, aproximadamente cada vez que completaban un mes de clases.

El módulo seguimiento del sistema, permitió acceder a encuestas aplicadas a los alumnos y con estos datos se prepararán informes y gráficas en las que se mostró

la satisfacción y el avance de los alumnos en sus materias. Los resultados se compartieron con los profesores de las asignaturas evaluadas y directivos, mediante cartas enviadas por correo electrónico.

Se espera que esta medida favorezca los procesos de reflexión y metacognición de los alumnos respecto a su desempeño global durante el semestre. A los maestros les facilitará la detección de elementos que estén impactando negativamente en sus clases, y con ello, tendrán la opción de ajustar sus estrategias y actuación docente de forma oportuna. A los directivos les será de utilidad para detectar casos particulares de las asignaturas que ameriten llevar a cabo acciones para profundizar en el análisis de alguna problemática y contribuir a su solución.

Para mantener un monitoreo y seguimiento de cualquier situación que pudiera presentarse durante el transcurso del semestre, en algunas sesiones síncronas de videoconferencia, cada dos o tres semanas, se dedicaba un tiempo para tratar las dudas de los estudiantes en la materia, dar orientación respecto a trámites o procesos importantes o comentar situaciones que pudieran afectar el desarrollo de las clases.

Para analizar aspectos de los estilos de aprendizaje dominantes en el grupo, a inicios del semestre se aplicó el cuestionario Honey-Alonso de estilos de aprendizaje (Alonso, Gallego y Honey, 1999), el cual mostró que existía en el grupo una preferencia dominante de aprendizaje Teórico y Pragmático. Las estrategias de enseñanza-aprendizaje se revisaron y ajustaron algunos detalles para coincidir con el estilo predominante del grupo, sin embargo, los alumnos con estilos de aprendizaje diferentes Activo o Reflexivo, se podrían ver afectados, por lo cual, también fue necesario diseñar actividades de otro tipo para considerar a estos estudiantes y, además, contribuir al desarrollo de otros estilos de aprendizaje en los demás alumnos.

### ***7.3.3 Diseño e Impartición de Curso para el Desarrollo de Habilidades Académicas***

El curso tuvo el propósito, por un lado, de proveer información relacionada con los procesos académicos y administrativos esenciales en la institución; además de brindar acompañamiento en el proceso de adaptación del alumno a la escuela, lo que se conoce como una modalidad de tutoría de asignatura, que sirve para orientar a los alumnos en su primer semestre escolar.

Por otro lado, el contenido del curso también aborda temas relacionados con el desarrollo de habilidades académicas para favorecer la motivación, las competencias digitales, la autorregulación y la metacognición. Las cuales son habilidades requeridas para afrontar la resolución de tareas o problemas, y que contribuyen de manera significativa en la adaptación de los estudiantes a su nuevo contexto educativo.

Para el diseño instruccional se siguió el modelo ADDIE y los contenidos y actividades del curso se implementaron en el ambiente virtual de aprendizaje (AVA), mediante la plataforma Moodle. El contenido se distribuye en cuatro unidades temáticas: 1) El entorno universitario, 2) Motivación como clave para el éxito, 3) Autodirección del aprendizaje, y 4) Mi visión como profesionista.

La planeación didáctica está disponible en el Anexo 2, en ella se muestra de forma puntual los contenidos, los recursos y momentos de su utilización. En la Figura 30 se presentan los principales ejes considerados en el diseño del curso, en donde, de manera transversal, el eje tecnológico sustenta el desarrollo de la propuesta.

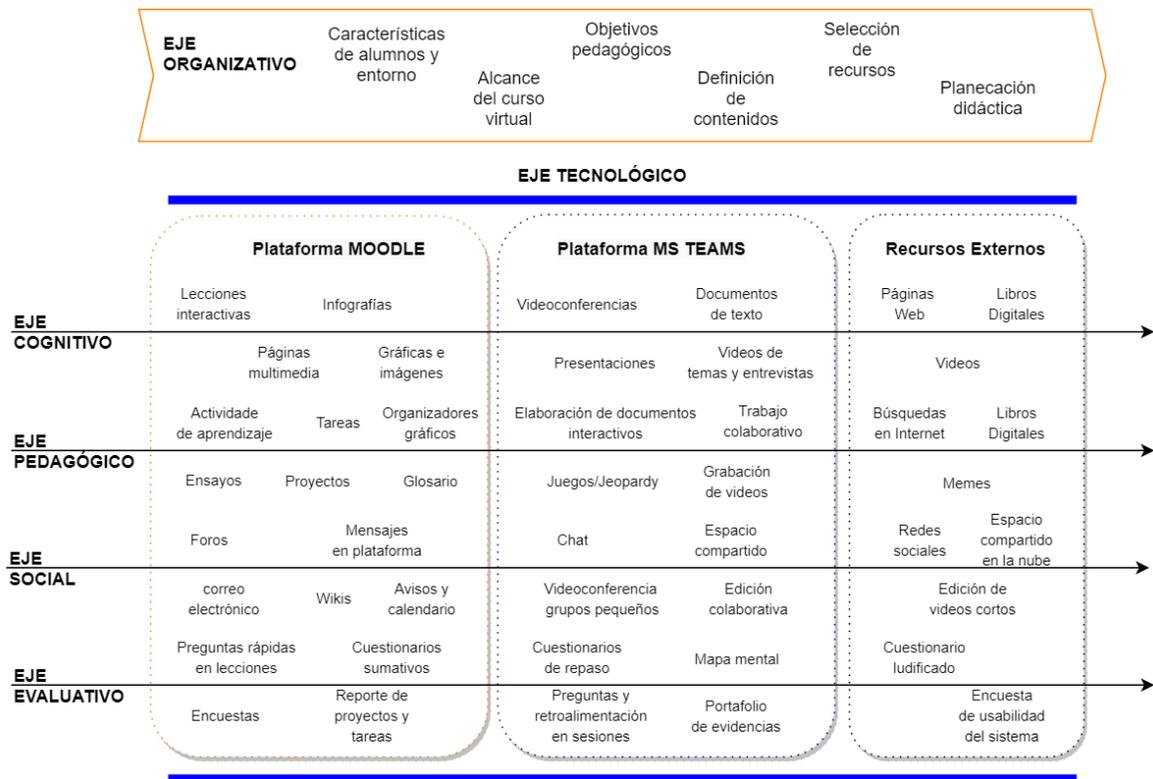


Figura 30 Elementos considerados en el diseño del curso.

Las actividades contempladas en el desarrollo del curso en la modalidad en línea fueron:

- Diseño y elaboración de recursos digitales con contenido multimedia: textos, páginas Web, videos, gráficas, imágenes, infografías, presentaciones, contenidos adaptativos, guiones de actividades.
- Diseño de actividades de aprendizaje mediadas con tecnología educativa para implementar estrategias que favorecieran el logro de los objetivos de aprendizaje (páginas interactivas, guiones de actividades, cuestionarios, lecciones multimedia, foros de discusión, elaboración de mapas mentales, elaboración de memes, actividades colaborativas, etc.)

- Implementación de los contenidos en la plataforma virtual, se emplearon herramientas disponibles en Moodle (lecciones, cuestionarios, encuestas, foros), en la plataforma Teams (bloc de notas de Teams, sesiones de videoconferencia, espacios compartidos de colaboración) y enlaces a recursos externos (calendarios, quizziz, genially).
- Impartición de sesiones de clase por videoconferencia en vivo, mediante la herramienta Teams. Además, se grabaron y se pusieron enlaces en el curso en Moodle para su visualización en modo asíncrono.
- Diseño de estrategias e instrumentos de evaluación, autoevaluación y coevaluación, con un enfoque centrado en el aprendizaje y el alumno. Se utilizaron cuestionarios de Moodle y Google Forms, elaboración de tareas en equipo y portafolios de evidencias.

En el Anexo 3, se presenta una tabla con el resumen de las características de los recursos tecnológicos empleados y la justificación de su empleo en el curso.

El uso de la plataforma Moodle y MS Teams ha permitido el desarrollo del curso a pesar de la pandemia, los alumnos han accedido a sesiones de videoconferencia en tiempo real, pero también han tenido la oportunidad de ver las grabaciones en cualquier momento que lo requirieron para repasar temas o recupera clases a las que no pudieron asistir por diversos motivos.

En general, el uso de las TIC como recursos didácticos aportan una gran flexibilidad para el diseño de actividades de aprendizaje, y dadas sus características intrínsecas, promovieron el estudio independiente, la repetición, la colaboración, la evaluación y retroalimentación de resultados, la presentación de contenidos en múltiples formatos, entre otras.

## **7.4 Análisis Cuantitativo de la Intervención**

El análisis cuantitativo se realiza a partir de datos obtenidos por encuestas diagnósticas y los datos históricos disponibles de las calificaciones de los alumnos. También, se analizan los resultados de la encuesta de satisfacción sobre sus asignaturas y su desempeño, realizadas de manera periódica durante el curso. Los instrumentos para medir y analizar el impacto de la intervención en aspectos cognitivos o de habilidades académicas, como la autodirección del aprendizaje, son importantes para determinar en qué medida las estrategias tuvieron éxito.

### ***7.4.1 Medición del Efecto de Intervención en el Perfil Autodirigido***

Como parte de la evaluación del plan de intervención, un aspecto clave fue la determinación de los efectos de la intervención académica en el desarrollo de habilidades cognitivas y de autodirección del aprendizaje.

La educación por medios virtuales requiere el despliegue de nuevas estrategias de enseñanza-aprendizaje. A su vez, implica que el estudiante es responsable de organizar y administrar su tiempo y esfuerzos, de manera que sean compatibles con su vida laboral, familiar y personal. Por tanto, su perfil de autodirección del aprendizaje debe verse fortalecido para lograr de manera satisfactoria sus metas educativas (Machado y Ochoa, 2017).

Los componentes de la autodirección o autorregulación del aprendizaje son múltiples, ya que incluye aspectos como la personalidad, la cognición, la motivación, los enfoques del aprendizaje y el contexto ambiental. Respecto a la medición de sus componentes, se han desarrollado diversos instrumentos sustentados en los diferentes enfoques y modelos con que se ha estudiado la autodirección (Aceves, 2008).

El Cuestionario de Indagación de Perfil Autodirigido Aumentado (CIPA+), propuesto originalmente por (Cázares, 2002), ampliado y validado por (Aceves, 2008); es un instrumento que se utiliza para la medición de la autodirección; consiste en una

prueba que permite definir el grado de su desarrollo mediante un análisis de cinco componentes.

Las autoras confirmaron la validez interna y la capacidad predictiva del CIPA+, el cual consta de 50 reactivos divididos en cinco componentes: a) planeación y selección de estrategias; b) autorregulación y motivación; c) independencia y autonomía; d) uso de la experiencia y conciencia crítica; y e) interdependencia y valor social. Además, la prueba determina un valor integrado de los cinco componentes en una escala inversa que se determinan como bajo, insuficiente, moderado, muy bueno y óptimo.

Metodología. El estudio tiene un enfoque cuantitativo de tipo cuasi-experimental con grupo de control no equivalente y mediciones antes y después (Hernández et al., 2014). Se busca determinar el efecto de una intervención educativa en las habilidades de autodirección de un grupo de estudiantes, a través del resultado global de la prueba CIPA+ y del análisis de los componentes de la autodirección. Los hallazgos permiten identificar los componentes más desarrollados y contrastar con un grupo de control para validar la pertinencia de la intervención.

La muestra está constituida por 56 alumnos de primer ingreso de carreras del área de computación, en la Facultad de Matemáticas de la Universidad Autónoma de Yucatán. En el Grupo 1 (sujeto a la intervención) participaron 28 alumnos de la carrera de Ciencias de la Computación; en el Grupo 2 (de control), 28 alumnos pertenecientes a la carrera de Ingeniería de Software accedieron a contestar el instrumento. La moda de la edad fue de 18 años en ambos grupos. En el Grupo 1 la media de edad fue de  $x=19.60$  y  $DS=2.6$ , con 4 mujeres y 24 hombres. Para el Grupo 2, la media de edad fue de  $x=19.07$  y  $DS=1.41$ , con 3 mujeres y 25 hombres.

El cuestionario fue diseñado en Google Forms y se compartió mediante la herramienta Teams, además del correo electrónico. La primera aplicación de la encuesta se realizó del 5 al 10 de octubre de 2020, al inicio de periodo escolar; la segunda aplicación del 1 al 6 de febrero de 2021, al finalizar las clases.

Las respuestas fueron almacenadas en una hoja de cálculo Excel, donde se realizaron las sumas de los puntajes de los diferentes componentes del instrumento. Posteriormente, los datos fueron importados y analizados desde el software SPSS versión 23.

#### **7.4.2 Resultados de la Encuesta CIPA+**

**Perfil de Autodirección por Niveles.** La Tabla 11 muestra el estado inicial del perfil de autodirección de forma global por niveles, antes de la intervención educativa en ambos grupos.

Tabla 11 Distribución de frecuencias por nivel de autodirección integrado, antes de la intervención.

NIVEL	GRUPO 1		GRUPO 2	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Óptimo	3	10.7%	4	14.3%
Muy bueno	5	17.9%	5	17.9%
Moderado	3	10.7%	7	25.0%
Insuficiente	4	14.3%	4	14.3%
Bajo	13	46.4%	8	28.6%

En el Grupo 1, el 60.7% de los alumnos se encuentran en niveles bajo e insuficiente y el resto, el 39.3%, se ubica en niveles moderado, muy bueno y óptimo, que dan cuenta de la existencia de habilidades de autodirección. En el Grupo 2, el 42.9% de los alumnos se ubican en niveles bajo e insuficiente, mientras que el 57.1% tiene un nivel que evidencia algún grado de las habilidades. Estas diferencias indican que, al inicio del semestre, un mayor número de alumnos del Grupo 2 tenían más desarrolladas las habilidades de autodirección.

La Tabla 12 muestra el estado final del perfil global de autodirección por niveles, una vez concluido el período lectivo para ambos grupos.

Tabla 12 Distribución de frecuencias por nivel de autodirección integrado, después de la intervención.

NIVEL	GRUPO 1		GRUPO 2	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Óptimo	7	25.0%	6	21.4%
Muy bueno	5	17.9%	3	10.7%
Moderado	4	14.3%	6	21.4%
Insuficiente	2	7.1%	5	17.9%
Bajo	10	35.7%	8	28.6%

En el caso del Grupo 1, sus integrantes estuvieron sometidos a una intervención educativa consistente en un curso de desarrollo de habilidades académicas. Después de la intervención, en el Grupo 1, se redujo a 42.8% la cantidad de alumnos en niveles bajo e insuficiente; el 25.0% de estudiantes está en el nivel óptimo. En el Grupo 2, el 46.5% de los alumnos quedó en niveles Bajo e Insuficiente, un porcentaje muy similar al del inicio del semestre.

**Perfil de Autodirección por Componentes.** En la Tabla 13 se muestran los resultados del Grupo 1 intervenido, derivados de la aplicación del instrumento CIPA+, al inicio y al final del curso con la intervención educativa implementada. Se indican los niveles para cada uno de los componentes de la autodirección y los estadísticos para determinar el nivel de significancia. El nivel de confianza es de 95%.

Los resultados en el Grupo 1 muestran mejoras en los niveles de cada uno de los componentes del perfil autodirigido y del resultado global del grupo. También se detecta que los componentes 4 y 5 son los únicos con diferencias estadísticamente significativas.

Tabla 13 Resultados de niveles de autodirección por componente, Grupo 1.

COMPONENTE	ANTES		DESPUÉS		t	p
	Media ±DS	Nivel	Media ±DS	Nivel		
1. Planeación y selección de estrategias	21.36 ±7.1	Insuficiente	18.29 ±5.8	Moderado	1.768	.083
2. Autorregulación y motivación	20.64 ±6.7	Insuficiente	18.14 ±5.5	Moderado	1.522	.134
3. Independencia y Autonomía	27.32 ±8.9	Bajo	24.82 ±7.3	Insuficiente	1.142	.258
4. Uso de la experiencia y la conciencia crítica	15.71 ±4.6	Moderado	13.50 ±3.2	Muy bueno	2.073	<b>.043</b>
5. Interdependencia y valor social	10.86 ±4.8	Muy bueno	9.07 ±2.0	Muy bueno	2.029	<b>.047</b>
Resultado global	95.89 ±27.7	Bajo	83.82 ±20.9	Moderado	1.840	.071

La Tabla 14 muestra los resultados del Grupo 2 de control, por componentes de la autodirección, al inicio y al final del curso. En este caso se trata de un grupo de control, en el cual los participantes no tomaron el curso con la intervención educativa. El nivel de confianza es de 95%.

Los resultados para el Grupo 2 no muestran mejoras importantes en los niveles de los componentes del perfil autodirigido, mejorando sólo un nivel en el resultado global, aunque aún es Insuficiente. Además, ninguno de los valores de los componentes tiene diferencias estadísticamente significativas, comparando el estado inicial y final del grupo.

Tabla 14 Resultados de niveles de autodirección por componente, Grupo 2.

COMPONENTE	ANTES		DESPUÉS		t	p
	Media ±DS	Nivel	Media ±DS	Nivel		
1. Planeación y selección de estrategias	20.64 ±8.2	Insuficiente	20.79 ±7.4	Insuficiente	-.068	.946
2. Autorregulación y motivación	19.86 ±7.2	Insuficiente	19.04 ±7.5	Insuficiente	.417	.678
3. Independencia y Autonomía	25.96 ±10.9	Insuficiente	26.50 ±10.5	Bajo	-.187	.852
4. Uso de la experiencia y la conciencia crítica	16.14 ±6.7	Insuficiente	15.71 ±6.5	Moderado	.241	.810
5. Interdependencia y valor social	11.00 ±5.2	Moderado	10.32 ±4.2	Muy bueno	.537	.593
Resultado global	93.61 ±36.6	Bajo	92.36 ±33.9	Insuficiente	.133	.895

A manera de resumen y discusión de los resultados obtenidos en esta etapa, respecto a los niveles de autodirección integrados (Tabla 11 y Tabla 12), podemos ver que, aunque inicialmente un mayor número de alumnos del Grupo 2 presentó un mejor perfil de autodirección, al finalizar el semestre en el Grupo 1 se redujo en mayor medida la cantidad de alumnos con carencias importantes en habilidades de autodirección. Se pasó de un 60.7% a un 42.8%, en el caso de los niveles Bajo e Insuficiente, y a la vez se incrementó de un 10.7% a un 25% el número de alumnos en nivel Óptimo. Lo anterior da cuenta de un impacto positivo de la estrategia de intervención desarrollada.

Al revisar los resultados derivados del análisis por componentes de la autodirección (Tabla 13 y Tabla 14), para ambos grupos se presentan niveles iniciales bajos, lo

cual ratifica la necesidad de intervenir y dar seguimiento a los grupos para madurar algunos de los componentes de la autodirección y de habilidades académicas, en este sentido la intervención desarrollada intenta incidir en tales aspectos.

El Grupo 1, como se observa en la Tabla 13, presenta un nivel global inicial Bajo, sin embargo, después de la intervención se ubicó en Moderado, subiendo dos niveles en el indicador general de la autodirección, con lo que se logró que el grupo evidenciara algún nivel de desarrollo en el perfil autodirigido. Sin embargo, los valores estadísticos ( $t=1.840$ ,  $p=0.071$ ) no representan una diferencia significativa a nivel global. Al verificar los valores de los componentes de la autodirección, se detectan mejoras en cada uno de ellos, pero sólo los componentes 4 y 5 muestran diferencias estadísticamente significativas ( $t=2.073$ ,  $p=0.043$ ) y ( $t=2.029$ ,  $p=0.047$ ), respectivamente.

Al analizar la Tabla 14 con los resultados del Grupo 2, estadísticamente no hay diferencias significativas en los valores obtenidos. Por tanto, después de la revisión de los resultados de los análisis por niveles y por componentes, se documentan resultados favorables para el Grupo 1, el cual fue sometido al proceso de intervención.

## **7.5 Análisis Cualitativo de la Intervención**

Para profundizar en la evaluación del plan de intervención, fue necesario analizar con un enfoque cualitativo las estrategias empleadas y los resultados alcanzados. Para ello, se partió de la pregunta de investigación cualitativa: ¿Qué estrategias de intervención educativa son apropiadas para modificar, mejorar el rendimiento escolar de alumnos con antecedentes problemáticos en la Facultad de Matemáticas de la UADY?

El diseño de las estrategias fue guiado por los resultados de la etapa diagnóstica y los conocimientos sintetizados en los modelos predictivos generados en la etapa de

la minería de datos educativos y uso del sistema predictivo. Surgieron algunos elementos necesarios en la definición de la intervención, por ejemplo, atender aspectos de hábitos de estudio, temas académicos particulares, habilidades académicas, metas profesionales, situaciones personales, motivación, antecedentes académicos, escuela de procedencia, conocimiento del modelo educativo, entre otros.

Para acotar la dimensión del análisis y poder generar códigos con un mayor nivel de precisión, a partir de la pregunta general, se elaboraron preguntas más específicas para identificar los aspectos puntuales relacionados con el propósito de la investigación:

PI 1: ¿Qué tipo de actividades fueron más significativas para el aprendizaje?

PI 2: ¿Qué tipos de conocimientos, habilidades o actitudes fueron favorecidos?

A partir de estos criterios se analizaron los datos provenientes de cuestionarios, observaciones de los profesores y documentos generados por los estudiantes, particularmente de los portafolios de evidencias solicitados para conocer la percepción de los alumnos sobre las actividades realizadas durante el curso, así como para analizar sus reflexiones respecto a los procesos y resultados obtenidos.

### ***7.5.1 Observaciones del Personal Académico Involucrado***

Derivados del desarrollo de las estrategias implementadas y la compartición de resultados preliminares, se recuperaron comentarios y mensajes recibidos por parte la secretaria académica, así como de los profesores y tutores de los alumnos. De igual forma, se registraron las observaciones de los dos profesores responsables de la impartición del curso.

### **Actividades o estrategias más significativas:**

- El envío de resultados de un posible riesgo académico de los alumnos, obtenidos a partir de las predicciones generadas por el sistema, hizo que varios de los participantes buscarán ayuda o más información, y estuvieran más al pendiente de sus clases. Algunos casos fueron turnados a los tutores para recibir consejo académico y orientación.
- La información sobre la percepción del desarrollo de los cursos, generada mediante la aplicación de encuestas periódicas y el uso del sistema predictivo, módulo de seguimiento, se consideró de utilidad pues proporcionó elementos para el rediseño o la mejora de las clases. El conocimiento de los aspectos analizados hizo que los profesores indagaran cómo mejorar sus métodos y estrategias de enseñanza-aprendizaje, así como temas relacionados con la motivación y la mejora de la comunicación. Se hicieron algunas sugerencias para mejorar la redacción de las preguntas y proporcionar más información para interpretar los resultados.
- La impartición del curso en modalidad e-learning y la grabación de las sesiones facilitó a los alumnos el acceso a las clases y los contenidos de los temas, desde cualquier ubicación y en cualquier momento, solventando de esta forma la problemática de la imposibilidad de clases presenciales.
- El uso de varias estrategias de comunicación y colaboración, por ejemplo, correo electrónico, foros, chat, video conferencias, contribuyó al desarrollo de habilidades sociales y a la integración del grupo. Las ideas y conceptos expresados por los estudiantes en sus participaciones en foros y actividades colaborativas dieron muestra del dominio de los temas o en su caso indicaron la necesidad de reforzar con más información o con otro tipo de actividades.

### **Aspectos favorecidos:**

- La información y conocimiento obtenidos fueron de gran utilidad para considerar nuevas estrategias de enseñanza, al identificar elementos

puntuales que requerían ser atendidos, por ejemplo, la motivación, por lo que se plantearon actividades para que los estudiantes reflexionarán al respecto y se fomentara su motivación intrínseca.

- Para la coordinación de la carrera y la secretaría académica, la información oportuna sobre el estado académico de los estudiantes y el desarrollo de sus clases posibilitó una mejor gestión de los aspectos administrativos de la educación.
- El realizar de forma regular actividades metacognitivas permitió a los estudiantes reflexionar sobre sus procesos de aprendizaje, autoevaluar su rendimiento, plantearse metas de formación y definir planes de actuación en consecuencia.
- El conocimiento de algunos aspectos emocionales y motivacionales contribuyeron a un mejor entendimiento de situaciones académicas y el fomento de habilidades para resolver problemas que partieron de un cambio actitudinal de los alumnos.
- El análisis de las competencias de su carrera y el conocimiento de información sobre el perfil de egreso, así como el conocer la opinión de maestros y egresados, contribuyó para que se visualizaran como profesionistas e identificaran las áreas de especialización.

### **7.5.2 Percepción de los Alumnos**

Con base en las preguntas, que definieron los criterios o categorías centrales de esta etapa de la investigación, se realizó el análisis de los datos contenidos en los portafolios de evidencias que presentaron los estudiantes al finalizar el curso. Se identificaron las unidades de análisis relevantes para dar paso a la definición de códigos que permitieran la categorización y el establecimiento de relaciones, con lo cual se tendrán los elementos para realizar inferencias y contrastar con la teoría.

El análisis de datos se realizó mediante el programa ATLAS.ti versión 9.1.2. Se recuperaron 9 documentos correspondientes a los portafolios de los estudiantes.

Cada portafolio contiene opiniones tanto personales como discutidas en equipo sobre varios aspectos del curso y sus resultados. Se analizaron los textos y los códigos se definieron conforme se fueron saturando con la información obtenida. Los segmentos de datos asociados a los códigos y las transcripciones se pueden consultar en el Anexo 4. En la Figura 31 se muestran los códigos obtenidos en el primer análisis de los documentos disponibles.



Figura 31 Nube de códigos del primer análisis.

**PI 1: ¿Qué tipo de actividades fueron más significativas para el aprendizaje?**

**Categoría Actividades Significativas.** En esta categoría se ubican los principales códigos que representan las actividades de aprendizaje desarrolladas por los estudiantes (ver Figura 32). Se presenta un resumen de las reflexiones que se hicieron, referentes a las razones por las cuales fueron elegidas.



Figura 32 Códigos de la categoría actividades significativas

Sub categoría áreas de interés. La actividad consistió en elaborar, de manera individual, un organizador gráfico en donde se sintetizarán las principales ramas o disciplinas de las ciencias de la computación. Luego, seleccionar y reflexionar sobre dos áreas de interés personal en las que quisiera especializarse el alumno, escribiendo las razones de su elección y que competencias habría que desarrollar para ser apto profesionalmente en ese campo.

Algunos de los aspectos que se destacaron sobre esta actividad fueron relativos al aprendizaje de las áreas que conforman las ciencias de la computación, ya que se desconocían. Consideran que tales conocimientos sirven para motivar, para fijar objetivos y definir planes de especialización, tener claras las expectativas de la industria, así como las oportunidades de trabajo, entre otros elementos.

Sub categoría competencias de egreso. La actividad se desarrolló de manera colaborativa para analizar las competencias de egreso para un Licenciado en Ciencias de la Computación. Se pidió que revisaran las competencias declaradas para su carrera, discutieran y seleccionaran las 5 más importantes explicando los motivos de su elección. También se pidió que analizaran algunos videos de egresados y profesionistas del área, preparados previamente.

En esta categoría se comentó que la actividad es de utilidad para conocer lo que se va a aprender durante la carrera y lo que se espera después de concluirla, es decir, orienta a los estudiantes sobre los aspectos centrales de su profesión. Sirvió para reflexionar si realmente se están desarrollando las competencias en las asignaturas y es un elemento que fomenta la motivación y la organización al tener claros los objetivos a los que se pretende llegar.

Sub categoría factores emocionales. La actividad se diseñó para trabajar en varios momentos. Momento individual: investigación sobre factores emocionales que pueden afectar el aprendizaje y el rendimiento académico. Lectura de material complementario y elaboración de un glosario de términos y conceptos investigados.

Momento en equipo: describir dos situaciones en las cuales el rendimiento académico se puede ver desbordado debido a un mal manejo de emociones. Explicación y propuestas de cómo la inteligencia emocional puede ayudar en tales situaciones. Reflejar los casos mediante la elaboración de tres memes por equipo. Momento social: compartición de los memes en la plataforma Teams y retroalimentación por parte de todo el grupo.

La actividad permitió a los estudiantes el conocer los factores emocionales que se deben de controlar y trabajar durante la vida universitaria. La inteligencia emocional es clave para canalizar los sentimientos positivos o negativos hacia el logro de metas de aprendizaje y mejorar en aspectos de rendimiento académico.

Sub categoría factores motivacionales. La actividad se desarrolló en varios momentos. Momento individual: lectura de material y reflexión individual sobre diferentes tipos de habilidades socioemocionales del aprendizaje. Momento en equipo: reflexión en equipo sobre las lecturas y respuesta a preguntas sobre la influencia de las habilidades emocionales en el aprendizaje. Se sintetizó el tema mediante la elaboración de tres memes por equipo. Momento social: compartición de los memes en la plataforma Teams y retroalimentación por parte de todo el grupo.

Con esta actividad se reflexionó sobre el papel de la motivación y su influencia en el estudio. Se consideró que la motivación es fundamental para perseverar en el logro de objetivos. El conocimiento de las distintas formas de motivación también sirve para buscar formas de mantenerse motivado y evitar que malos resultados, emociones negativas u otros aspectos afecten en el desempeño escolar.

**PI 2: ¿Qué tipos de conocimientos, habilidades o actitudes fueron favorecidas?**

El análisis de las opiniones de los alumnos, sobre aspectos favorecidos durante el curso, generó los siguientes códigos: autocontrol, competencias digitales, establecimiento de metas, estrategias de estudio, hábitos de estudio, manejo de

emociones, metacognición, motivación, planificación y tipos de aprendizaje. Las transcripciones de los datos resultantes se pueden consultar en el Anexo 4.

Se presenta un resumen de las opiniones los estudiantes en cada una de las categorías de códigos:

**Autocontrol.** El autocontrol hace referencia a comportamientos adquiridos por la persona y que se encaminan a modular o cambiar el propio comportamiento motivado fundamentalmente por procesos hedónicos (Navarro, 2003). Es decir, el esfuerzo consciente de la persona por modificar sus reacciones, ante situaciones positivas o negativas. Es parte de la inteligencia emocional necesaria para interactuar con el mundo teniendo en cuenta los sentimientos.

En esta categoría, los estudiantes aprendieron a identificar factores emocionales que les podían afectar, reflexionando sobre opciones para regularlos y poder tomar buenas decisiones. El conocimiento de sus sentimientos y reacciones fue importante para afrontar situaciones académicas o extraescolares que afectan su rendimiento académico.

**Competencias digitales.** Se define las competencias digitales como el uso creativo, crítico y seguro de las tecnologías de la información y las comunicaciones para alcanzar los objetivos relacionados con el trabajo, la empleabilidad, el aprendizaje, el tiempo libre, la inclusión y la participación en la sociedad (INTEF, 2017). Algunas de sus áreas son la información y la alfabetización informacional, la comunicación y la colaboración, la creación de contenidos digitales, la seguridad y la resolución de problemas.

Sobre esta categoría los participantes consideran que desarrollaron más habilidades y conocimientos acerca de los distintos tipos de plataformas y entornos virtuales, lo que les facilitó el aprendizaje y la comunicación en el entorno académico, además de acceder a variadas fuentes de información. De igual forma, utilizaron medios de comunicación y colaboración, lo que les permitió realizar

actividades a distancia, y reconocieron la importancia de la tecnología en la actualidad, particularmente en la educación.

**Establecimiento de metas.** Una meta refleja los propósitos del estudiante, implícitamente involucra la cantidad, calidad y nivel de desempeño. La perseverancia depende del establecimiento de metas y de las autoevaluaciones sobre su progreso. El establecimiento de metas significa establecer un objetivo que impulse la actuación. Las metas pueden ser establecidas por los propios aprendices o seguir las que otros les proporcionen, por ejemplo, profesores, compañeros, padres o profesionistas tomados como referentes (Schunk, 2012).

En este sentido, en la opinión de los alumnos, el conocer las áreas de la carrera y la realidad del ámbito profesional, incluyendo la opinión de profesores y egresados, fue un factor que les facilitó la definición de metas de aprendizaje e influyó en la motivación para continuar con sus estudios y prepararse mejor para el entorno laboral.

**Estrategias de estudio.** Tienen que ver con el crear un clima psicológico adecuado para el aprendizaje. Incluyen también estrategias para comprender, retener, recuperar y usar la información. Por ejemplo, tomar notas, escuchar, subrayar, resumir y afrontar el estrés; y estrategias para presentar exámenes y organizar un escenario para el aprendizaje, entre otras. Por el contrario, el no contar con esta habilidad puede ocasionar que el estudiante no complete sus tareas de manera eficiente (Farrington et al., 2012).

Algunos comentarios en este sentido indican un aprendizaje de estrategias de trabajo, individual y en equipo, para optimizar el tiempo y realizar el trabajo de manera más eficiente. Las estrategias empleadas permitieron una mejor comprensión y realización de las actividades.

**Metacognición.** La metacognición se refiere al control deliberado y consciente de las actividades mentales; incluye conocimientos y actividades de vigilancia diseñadas

para garantizar que las tareas se realicen con éxito. Se relaciona con procesos reflexivos que permiten la planificación, la comprensión y la vigilancia de las estrategias para lograr los objetivos de aprendizaje (Torras, 2019).

Los segmentos de información dan cuenta del proceso mediante el cual los alumnos reconocieron su entorno y los elementos que influían en su educación. Se reflexionó sobre los valores y aptitudes que se deben tener como estudiantes y egresados, así como los retos por afrontar y las competencias a desarrollar.

Manejo de emociones. Los pensamientos y sentimientos de las personas pueden ser positivos o negativos, lo que influye en la manera en que se motivan a sí mismas y perseveran durante las dificultades. La forma como lidian con las emociones, especialmente durante periodos de estrés, indica la capacidad para superar obstáculos y en tomar decisiones en momentos críticos (Benight y Bandura, como se citó en Schunk, 2012).

Los alumnos indicaron que aprendieron reconocer factores emocionales que intervienen en el proceso educativo, y como estos pueden ser canalizados hacia metas de aprendizaje mediante la motivación intrínseca y la inteligencia emocional. También se identificaron sentimientos negativos que causan la falta de motivación para estudiar, por ejemplo, el estrés causado por los exámenes o por sacar una mala nota, ante ello, la inteligencia emocional es fundamental para sobreponerse a este tipo de situaciones.

Motivación. La motivación es un proceso que influye para que una persona realice una acción o piense de modo que persista en actividades para el logro de metas. Hay un vínculo recíproco entre motivación y autorregulación, las personas motivadas realizan actividades autorregulatorias que les pueden servir de ayuda para alcanzar sus propósitos (Farrington et al., 2012).

Los estudiantes conocieron los tipos de motivaciones y su importancia en el aprendizaje, y como puede ser afectada por las emociones y el entorno. El

plantearse metas de desarrollo profesional generó motivación al visualizar las posibilidades de la carrera.

Planificación. La estrategia autorreguladora de planificación tiene que ver con el establecimiento de un plan de acción; el cual debe incluir la identificación o determinación de metas de aprendizaje, la predicción de resultados, además de la selección y programación de estrategias, incluye la administración del tiempo. Sirve para facilitar la ejecución de la tarea, incrementar la probabilidad de éxito y la calidad de la ejecución o del producto obtenido (Díaz Barriga y Hernández Rojas, 2010).

Los estudiantes consideran que los aprendizajes obtenidos en el curso fueron un punto de partida importante para definir y a organizar las actividades escolares de una manera responsable y así obtener un aprendizaje significativo. A su vez, les permitió visualizar a futuro para poder definir caminos que los preparen para lograr sus aspiraciones.

Hábitos de estudio. Son un conjunto de rutinas y técnicas que emplean habitualmente los estudiantes para adquirir y utilizar los conocimientos que adquieren. Están relacionados con las estrategias de estudio e influyen positivamente en el rendimiento escolar y donde intervienen factores como el interés y la motivación interna. Es deseable que la escuela fomente los hábitos que promuevan el máximo aprovechamiento de las capacidades de los estudiantes y contribuyan a neutralizar los efectos nocivos de un ambiente familiar o social desfavorable (Schunk, 2012).

Los estudiantes reconocieron los hábitos que tienden a tener los estudiantes, hábitos tanto positivos como negativos y que afectan el desempeño académico en la universidad. Las reflexiones y actividades ayudaron a una mejor administración del tiempo libre y del estudio, también se destacó el tiempo de descanso suficiente para rendir de una mejor manera.

Tipos de Aprendizaje. Se relaciona con los contenidos vistos en el curso sobre aspectos teóricos del aprendizaje, el conocimiento del proceso de aprendizaje aclaró dudas y proporcionó una base para analizar sus propios de formación. Al identificar las distintas formas y modalidades de estudio se pudieron seleccionar las estrategias de estudio que mejor se adaptaron a sus estilos y características personales.

## 8. INTEGRACIÓN DE RESULTADOS Y DISCUSIÓN

En este capítulo se presentan, de manera resumida y ordenada, los principales resultados de los procesos desarrollados durante las etapas de la investigación. A su vez, se contrastan y discuten los hallazgos de cada fase del estudio con algunos referentes y trabajos previos. También, se discuten las aportaciones de este trabajo, en cuanto a su enfoque integrador y holístico, y finalmente, se realiza la comprobación de la hipótesis y de los supuestos de la investigación.

### 8.1 Resultados de la Etapa de Minería de Datos

A partir del proceso de la minería de datos educativos se generó un modelo que describe los pasos para aplicar los algoritmos de predicción de riesgo académico (ver Figura 33) y que muestra las relaciones de las principales tareas a realizar en la implementación del sistema.

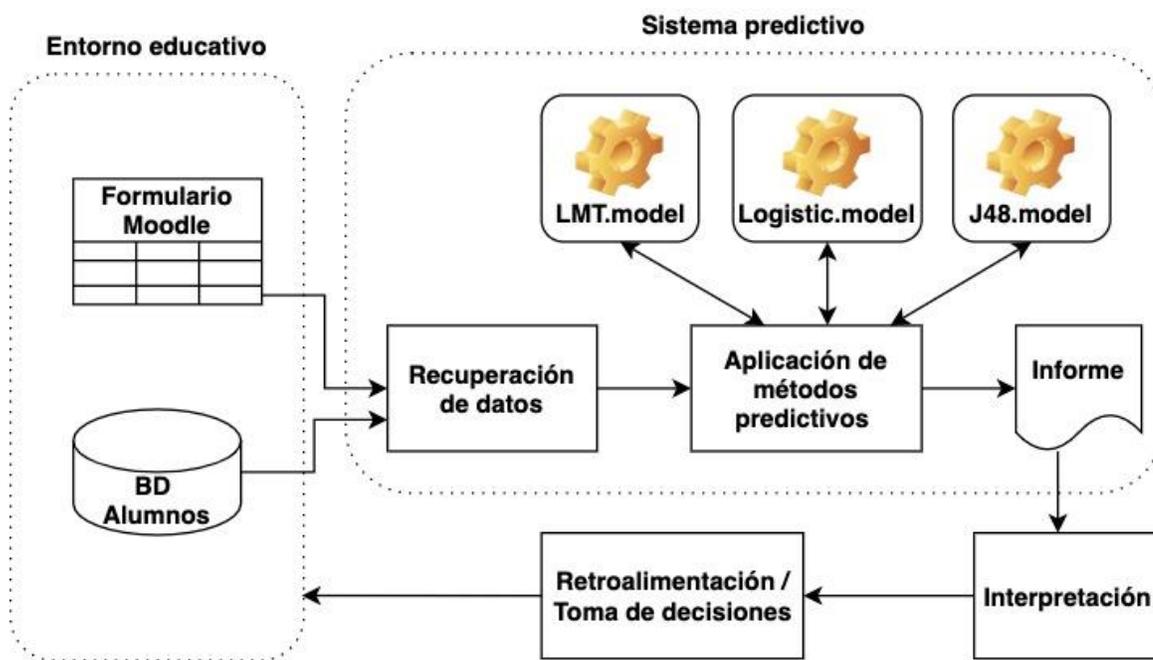


Figura 33 Modelo para el desarrollo y aplicación del sistema predictivo.

Para la generación de los modelos, fue necesario analizar cuáles de los atributos disponibles aportaban mayor poder predictivo y determinar cuáles algoritmos o combinaciones de algoritmos resultaban más convenientes para ser implementados en una herramienta de software. El subconjunto de variables predictoras se redujo a 9 atributos, lo cual facilitó y agilizó la recuperación de datos, ya sea que fueran solicitados directamente a los alumnos o que se recuperen desde algún repositorio digital.

Después de analizar los resultados de los algoritmos de clasificación, se eligieron los modelos LMT y Logistic para ser implementados en la aplicación de software; ya que, para la variable dependiente RIESGO\_P, obtuvieron los valores de precisión más altos y sus valores de AUC ROC los ubican en el nivel de Test bueno, según (Minguillón et al., 2017). Se incluyó además el modelo basado en el algoritmo J48 porque es de gran ayuda para entender las relaciones entre las variables predictoras.

También, como lo explican (Witten et al., 2011), mediante la herramienta WEKA se generaron archivos con los modelos producidos por los algoritmos de minería de datos. Lo que permitió que, mediante el sistema desarrollado con el lenguaje de programación Java, se pudieran recuperar estos modelos y crear instancias para realizar la clasificación y predicción de nuevos casos de alumnos.

De los modelos creados, en general, se vio que el resultado de la prueba de competencias disciplinares DIAG y la prueba diagnóstica de competencias básicas ICNE tuvieron el mayor poder predictivo y señalan en buena medida el bajo rendimiento escolar de los estudiantes; en menor medida también influyen: el resultado final del Exani-II, la carrera elegida, la preparatoria de origen, la edad y el responsable del alumno.

En relación con la carrera elegida, se pudo verificar que los grupos de la Licenciatura en Ciencias de la Computación tienden a presentar situaciones de riesgo académico de manera más recurrente, por lo que dan cuenta de la necesidad de atender

aspectos relacionados con los antecedentes académicos de los alumnos; y también señalan algunas situaciones de índole socioeconómicas de los estudiantes, por ejemplo, cuando el responsable del alumno no es el padre se presentan más casos de bajo rendimiento académico, lo que sugiere el despliegue de estrategias más efectivas de seguimiento y apoyo ante tales condiciones.

Como ha sido señalado en los trabajos citados en la sección de antecedentes, en la determinación del rendimiento académico convergen una gran cantidad de variables, no sólo de orden académico o del entorno escolar, sino del contexto sociodemográfico o de aspectos cognitivos e interpersonales, entre otros.

## **8.2 Resultados Generados por el Sistema Predictivo**

Considerando los elementos del modelo de desarrollo del sistema predictivo ilustrados en Figura 33, para la detección de alumnos en riesgo académico, se llevaron a cabo los pasos necesarios para su implementación. Una vez concluido, y al disponer de una herramienta de software con las características del sistema desarrollado, se pudo analizar de manera más rápida y efectiva una gran cantidad de datos relacionados con los antecedentes académicos y no académicos de los alumnos, y su empleo ayudó a determinar grupos y alumnos en posible riesgo académico.

### **8.2.1 Predicción**

Al utilizar el módulo de predicción del sistema, se seleccionó el modelo LMT con la clase riesgo basada en promedio (Riesgo\_P), pues su nivel de exactitud fue el más alto, con un 75.42% de precisión. Es decir, si se determina que existe riesgo académico, se está considerando que el alumno probablemente obtenga un promedio de calificaciones menor a 75 en el primer semestre de estudios. Este indicador de riesgo es más o menos equivalente a pronosticar que el alumno reprobará 2 o más materias, cómo lo sugiere el uso del modelo basado en la clase Riesgo\_A, basada en la cantidad de materias reprobadas.

El resumen de los pronósticos y los resultados reales obtenidos por los estudiantes se muestran en la matriz de confusión de la Tabla 15.

Tabla 15 Matriz de confusión de los resultados de la predicción.

	Predicción SI	Predicción NO	Total
Real SI	VP=8	FN=1	9
Real NO	FP=10	VN=8	18
Total	18	9	N=27

En donde:

VP (Verdaderos positivos): instancias positivas correctamente clasificadas por el sistema.

FN (Falsos negativos): instancias que son positivas y que el sistema dice que no lo son.

FP (Falsos positivos): instancias que son negativas pero el sistema dice que no lo son.

VN (Verdaderos negativos): instancias que son correctamente clasificadas como negativas.

N = es el número total instancias,  $N=VP+FN+FP+VN$ .

En la Tabla 16, se presentan las tasas de acierto esperadas según el modelo del algoritmo predictivo LMT y se contrastan con las tasas de acierto de las instancias reales correctamente clasificadas.

Tabla 16 Tasas de acierto de la predicción y resultados reales.

	Real	Predicción
Tasa de acierto	0.593	0.754
Tasa de error	0.407	0.246
Tasa de acierto SI	0.444	0.693
Tasa de acierto NO	0.889	0.805

En donde:

Tasa de acierto:  $VP+VN/N$

Tasa de error:  $FP+FN/N$

Tasa de acierto SI:  $VP/VP+FP$

Tasa de acierto NO:  $VN/VN+FN$

Al confrontar los valores reales con los esperados, se ve que la tasa de aciertos para el valor Real SI (0.444) es menor a la pronosticada (0.693), y para el caso del valor NO, la tasa de acierto (0.889) es mayor que la esperada (0.805). Estas diferencias se explican por el proceso de intervención educativa al que fue sujeto el grupo que aportó las instancias de los casos reales. Es decir, se presenta una disminución de casos de alumnos en riesgo académico.

### **8.2.2 Seguimiento**

El módulo de seguimiento del sistema permitió analizar las encuestas de opinión respecto a: Q1) la pertinencia de los contenidos, Q2) el método o estrategias de aprendizaje, Q3) la motivación en el desarrollo de las actividades, Q4) el desempeño como alumno, y Q5) la comunicación con el profesor. También, se obtuvo un indicador de satisfacción global de cada curso, el cual se calculó promediando los valores todos los ítems de encuesta.

El instrumento se aplicó en tres ocasiones durante el semestre y los resultados globales se muestran en la Figura 34, en donde se observan las tendencias en la satisfacción de los cursos.

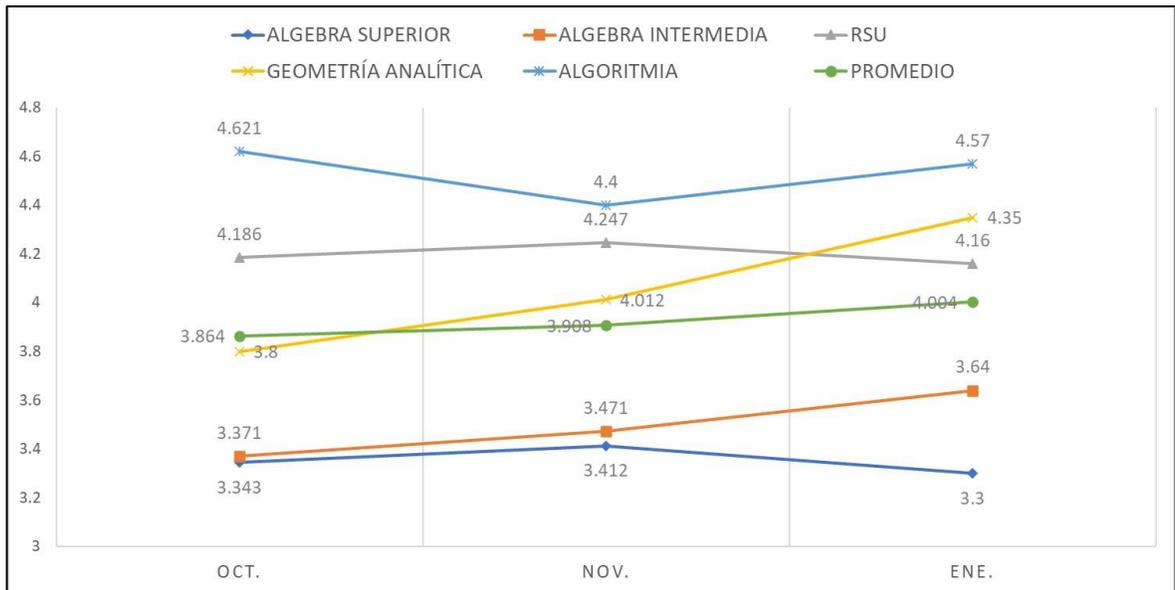


Figura 34 Resultados globales de satisfacción de los cursos.

Las materias algoritmia y RSU se mantuvieron en un buen nivel de aceptación por parte de los alumnos, entre bueno y excelente. Geometría analítica, comenzó con un nivel regular y al finalizar el curso se evaluó como bueno. Algebra superior y algebra intermedia se mantuvieron siempre en un nivel regular, por debajo del promedio de todas las asignaturas.

Los resultados de las encuestas se enviaron a los profesores y algunos de ellos expresaron su aprobación por la información recibida, ya que les permitiría estar al tanto de la opinión de los alumnos y hacer ajustes en caso de ser necesario.

Para la asignatura álgebra superior, dado que fue la peor evaluada, se presentan las tres gráficas generadas por el sistema, en las que se muestran los valores para los ítems de opinión, con el propósito de identificar los elementos que influyeron en su nivel obtenido (ver Figura 35).

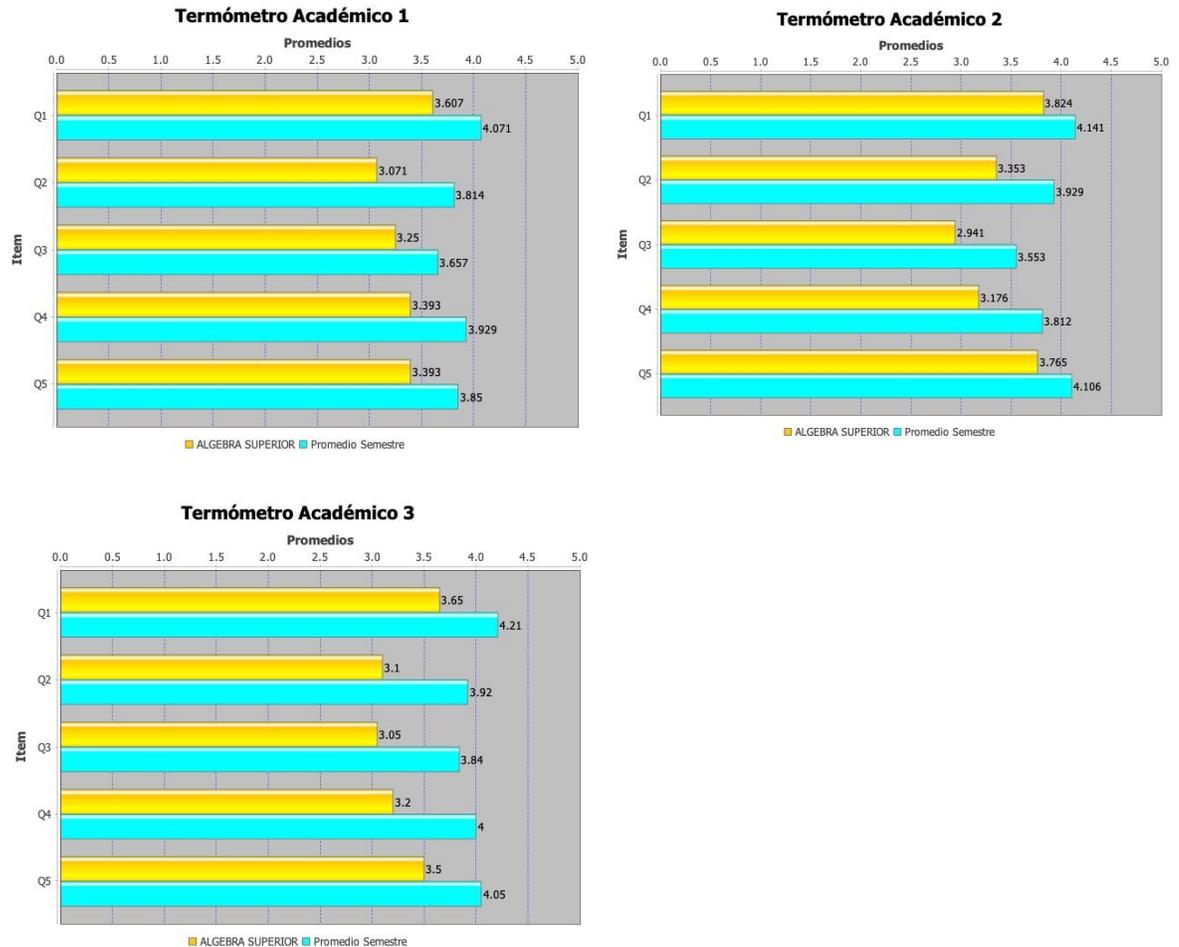


Figura 35 Gráficas de resultados de satisfacción para Álgebra superior.

Los elementos peor evaluados fueron: Q2) el método o estrategias de aprendizaje, Q3) la motivación en el desarrollo de las actividades y Q4) el desempeño como alumno. Se recuerda que los cursos fueron impartidos en modalidad virtual, y posiblemente las materias más afectadas por las limitaciones en cuanto a su forma de impartición fueron las del área de matemáticas, que comúnmente requieren del

uso de pizarrones para desarrollar ejemplos y ejercicios de manera adecuada. En cambio, en algoritmia se pueden utilizar herramientas computacionales (compiladores, intérpretes, editores de diagramas de flujo) y TIC de una forma más natural y sencilla. Para los estudiantes, el ejercicio de reflexionar sobre el desarrollo de las asignaturas de manera regular contribuyó al desarrollo de habilidades metacognitivas y a la autodirección de su aprendizaje

### **8.3 Resultados del Diagnóstico y la Planeación de la Intervención**

En esta fase del proceso de intervención se lograron identificar varios factores personales del alumno y del contexto que afectan su rendimiento, por ejemplo, grupos de alumnos con pobre rendimiento en los exámenes de ingreso, falta de habilidades cognitivas en ciertas áreas de estudio, debilidades en habilidades metacognitivas y de autodirección, como la motivación, el establecimiento de metas y la organización de sus tiempos.

También, relacionados con su contexto, hay elementos que influyen en los estudiantes de nuevo ingreso, como la falta de apoyo tutorial y orientación efectiva en su primer año de estudios, además de la necesidad de un mayor seguimiento e interacción con los estudiantes, por parte de tutores, asesores y directivos. Estos elementos sirvieron para determinar las estrategias de intervención a implementar en las siguientes etapas.

Para la intervención, se logró el diseño, planificación e implementación de las actividades del curso en línea Habilidades académicas para alumnos universitarios de primer ingreso (Anexo 2). El curso se ofrecerá de manera regular para las nuevas generaciones de alumnos que lo necesiten. El curso fue implementado en la plataforma de gestión de aprendizaje Moodle. Su estructura y organización facilita la navegación y el acceso a lecciones y páginas para manejar los contenidos de los temas del programa del curso; se desarrollaron e incluyeron recursos digitales, como videos, infografías, páginas interactivas, presentaciones, y diversas actividades de aprendizaje para lograr el aprendizaje significativo de los alumnos.

## 8.4 Resultados de Aspectos Cuantitativos de la Intervención

Se presentan los resultados obtenidos en cuanto a datos cuantitativos, para ello se consideran las calificaciones históricas de los alumnos y las calificaciones obtenidas durante el semestre en que se realizó la intervención, con lo cual se puede determinar su impacto. De igual forma, para conocer la influencia en la autodirección se contrastan los valores del perfil autodirigido antes y después de la intervención.

### 8.4.1 Efectos en el Rendimiento Académico

Los parámetros utilizados para determinar el rendimiento académico fueron los promedios de calificaciones y la cantidad de materias aprobadas en el semestre. En la Figura 36, se presentan los porcentajes de alumnos que obtuvieron calificaciones en los rangos indicados. Considerando los alumnos que obtuvieron un promedio entre 0 y 69, se evidencia una mejora al reducir de un 57%, en resultados históricos de grupos LCC (2016-2018) a un 22%, en el grupo intervenido (2020).

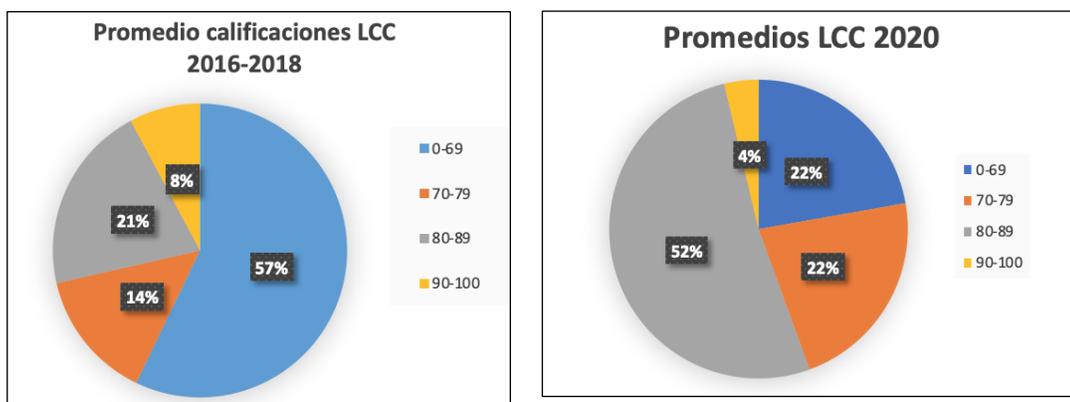


Figura 36 Comparación de promedios de calificaciones, históricas y grupo intervenido.

Al revisar la cantidad de alumnos con materias reprobadas, como se muestra en la Figura 37, en el grupo intervenido sólo el 26% de los alumnos tuvo 2 o más materias reprobadas al finalizar el semestre, mientras que en los resultados históricos (2016-2018) el 64% reprobaba dos o más asignaturas.

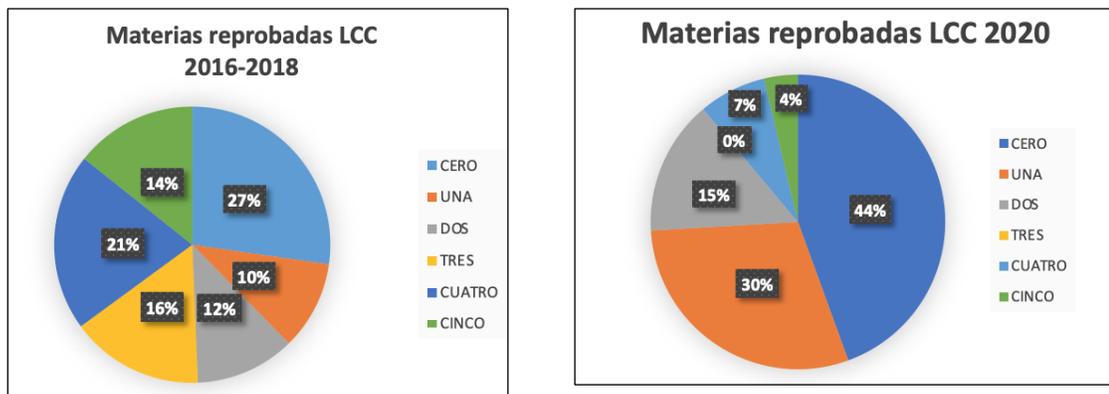


Figura 37 Comparación de materias reprobadas, históricas y grupo intervenido.

Tanto en materias reprobadas como en promedios de calificaciones, se evidencia una mejora en el rendimiento académico. Sin embargo, se debe de tomar con reserva este resultado ya que la impartición de las clases del grupo sujeto a la intervención fue en modalidad virtual o en línea, debido a las condiciones generadas por la pandemia; mientras que los resultados históricos corresponden a cursos impartidos en modalidad presencial.

#### **8.4.2 Efectos en el Perfil Autodirigido**

La Figura 38 muestra los resultados del perfil de la autodirección según la frecuencia de los niveles obtenidos por los grupos, antes y después de la intervención. Inicialmente el 60.7% de los alumnos de grupo 1 se encontraba en niveles Bajo o Insuficiente, mientras que el 42.9% de los alumnos del grupo 2 estaba en los mismos niveles. Después del curso el 42.8% de los alumnos de grupo 1 se ubicó en niveles Bajo o Insuficiente; sin embargo, en el grupo 2 el 46.8% de los alumnos se percibieron con niveles Bajo e Insuficiente.

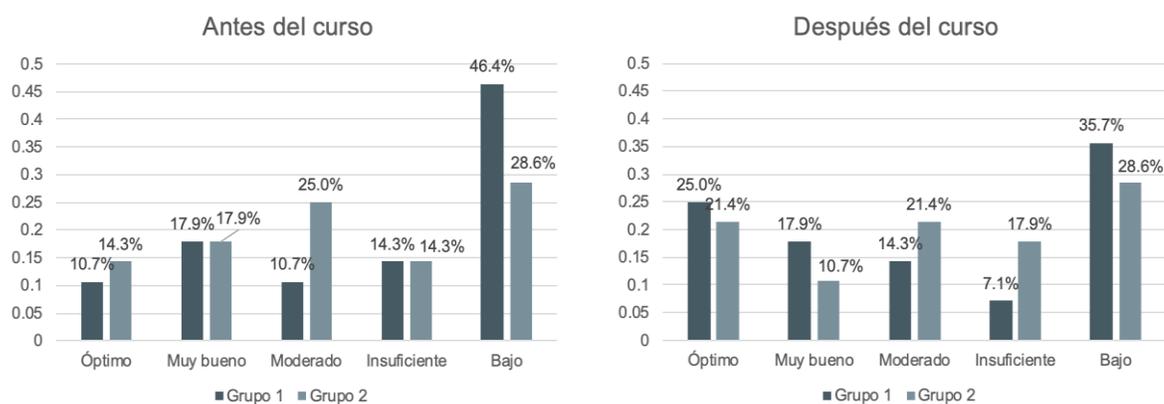


Figura 38 Distribución de frecuencias por niveles de autodirección, antes y después del curso.

Los resultados de perfil de autodirección por componentes, en el primer momento, mostraron que el grupo 1 tenía niveles Bajo e Insuficiente en los componentes de Planeación y selección de estrategias; Autorregulación y motivación e Independencia y autonomía. Sólo para los componentes de Uso de la experiencia y conciencia crítica e Interdependencia y valor social, se mostraron niveles Moderado o Muy Bueno; lo que ubicó al grupo en un nivel global Bajo.

Se destaca la utilidad del instrumento para identificar grupos que requieren procesos de intervención, al detectarse niveles bajos o insuficientes en sus perfiles de autodirección. Los resultados en la identificación de grupos en riesgo son consistentes con los obtenidos a partir de modelos predictivos de riesgo académico empleando técnicas de minerías de datos educativos, pues también mostraron que algunos programas de estudios tienen una mayor propensión al fracaso escolar por sus antecedentes académicos, lo que ratifica la necesidad de fortalecer algunas de sus habilidades.

Al aplicar por segunda vez la encuesta y contrastar, se ha podido evidenciar el impacto que tuvo la intervención educativa. Sólo el componente de Independencia y autonomía obtuvo un nivel Insuficiente, lo demás componentes se ubicaron en niveles Moderado o Muy Bueno, con un resultado global Moderado. Los

componentes que mostraron diferencias estadísticamente significativas después de la intervención fueron: componente 4, Uso de la experiencia y la conciencia crítica; y componente 5, Interdependencia y valor social.

En términos generales, los resultados del cuestionario CIPA+ muestran mejoras en los niveles de autodirección del grupo sujeto a la intervención, al pasar de un nivel Bajo a Moderado. También, el grupo evidencia un incremento en alumnos con algún nivel de autodirección (inicial 39.3%, final 57.2%). El perfil de la autodirección tuvo un cambio con un efecto moderado ( $r=0.333$ ).

En el grupo de control el nivel paso de Bajo a Insuficiente, sin que ninguno de sus componentes obtuviera cambios significativos estadísticamente. Ambos grupos tomaron sus clases en modalidad virtual o en línea, es decir en las mismas condiciones, por lo que los hallazgos aportan mayor grado de certidumbre en cuanto a la eficacia de las estrategias de intervención desplegadas.

En relación con la estrategia de intervención basada en e-learning, al emplear recursos basados en entornos virtuales de aprendizaje, los estudiantes tuvieron la oportunidad de explorar diferentes tipos de contenidos y estrategias didácticas, favoreciendo su aprendizaje y el desarrollo de habilidades académicas, lo cual es consistente con los hallazgos aportados por Elcullada, Galang y Hallar (2021).

Aspectos de la autodirección como la interdependencia y valor social, relacionados con la valoración de la diversidad de las personas, se vieron favorecidos al realizar actividades en equipo basadas en herramientas de comunicación y colaboración; además de impactar positivamente en la motivación, como también señalan (Benítez et al., 2014), al emplear las TIC en proceso educativos.

### **8.5 Resultados de Aspectos Cualitativos de la Intervención**

De las observaciones realizadas por el personal académico, se destacan varios aspectos que fueron favorecidos por las diversas estrategias desplegadas en la intervención. En la gestión y práctica educativa se mejoró en el diseño de clases, la

detección de situaciones de riesgo y la atención de aspectos específicos requeridos por los alumnos, como una mejor comunicación y considerar la motivación como elemento importante del desarrollo de los cursos.

Los instrumentos y sistemas desarrollados contribuyeron a dar un mejor seguimiento de los cursos, generando información para la toma de decisiones oportuna y adecuada a las necesidades detectadas. También, para los alumnos, las actividades de análisis y reflexión fomentaron sus habilidades metacognitivas. El diseño del curso en modalidad e-learning con variedad de actividades interactivas y de trabajo colaborativo, contribuyó al desarrollo de sus competencias digitales, además de desarrollar habilidades de comunicación y trabajo en equipo. Algunas actividades permitieron una reflexión inicial de elementos motivacionales, emocionales y de visualización a futuro, que les permitiera plantearse metas de crecimiento académico y profesional.

En cuanto a la opinión de los alumnos para responder a la pregunta sobre los aprendizajes favorecidos, en el mapa semántico de Figura 39 se interrelacionan los diferentes conceptos extraídos a partir del análisis de los portafolios de evidencias.

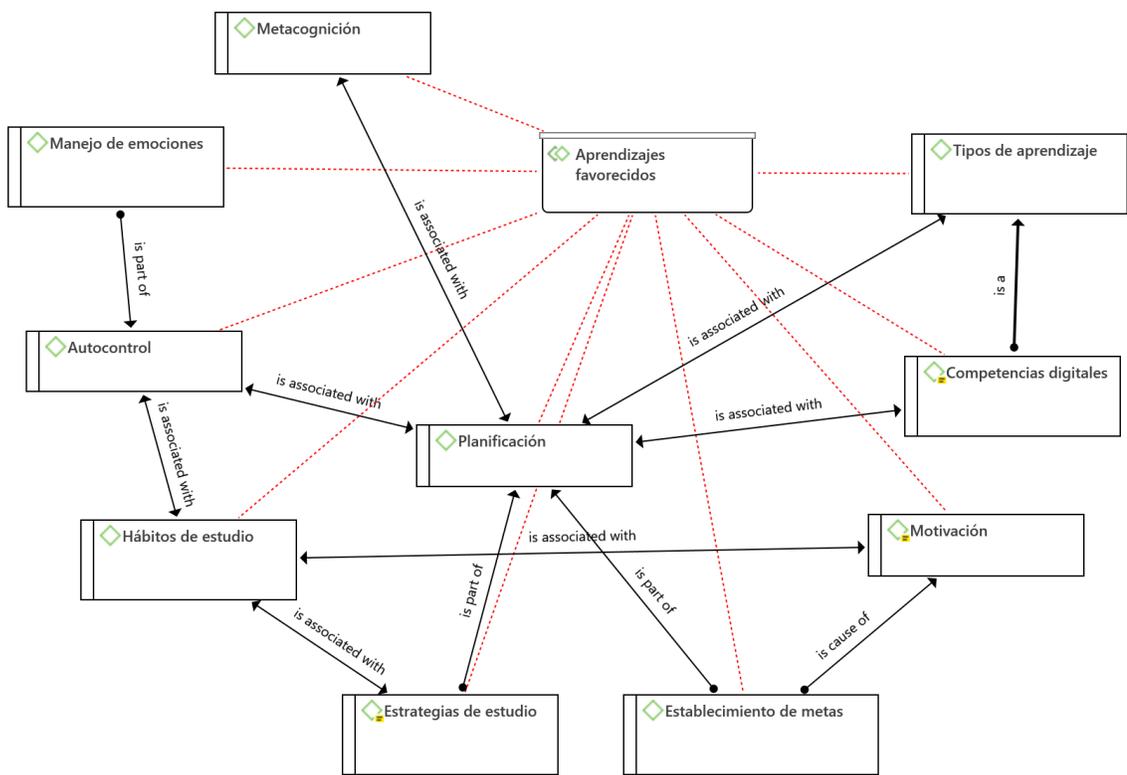


Figura 39 Red semántica para la categoría aprendizajes favorecidos.

Se infieren una serie de relaciones entre los elementos considerados como aprendizajes favorecidos. Se observa una relación entre el manejo de emociones como elemento requerido para lograr el autocontrol. El establecimiento de metas es una de las causas que proporcionan motivación, que su vez se relaciona con los hábitos de estudio. Las competencias digitales son un tipo de aprendizaje y a su vez están asociadas con el desarrollo de los contenidos y actividades definidas en la planeación, más aún en un curso desarrollado en línea. También, el establecimiento de metas y la definición de estrategias de estudio son parte fundamental para considerar en la planeación. El autocontrol está relacionado con la definición y seguimiento de la planificación. Finalmente, la metacognición está asociada con la planificación, pues proporciona elementos para analizar su diseño, desarrollo, seguimiento o actualización.

Los hallazgos del análisis cualitativo están en consonancia con los referentes teóricos, y también se observan dos procesos o habilidades que se interrelacionan estrechamente para el logro efectivo del aprendizaje.

Por un lado Chaves et al. (2015), indican que las habilidades metacognitivas dotan a las personas de la capacidad de reflexionar sobre lo que necesitan para aprender y sus logros cognitivos; también, les permite articular lo aprendido con experiencias y conocimientos previos que les ayuda en la solución de nuevas problemáticas. Es el control consciente y deliberado de la actividad cognitiva.

Por otra parte, como señala Schunk (2012), la autoregulación del aprendizaje es fundamental en la educación ya que permite a los aprendices regular la tarea por realizar, definir metas y planes de actuación, así como tácticas de estudio, y adaptarse antes situaciones cambiantes o adversas.

Los resultados también abonan en el plano del rendimiento mediato, como es explicado por York et al. (2015) en su modelo conceptual de éxito académico, ya que además del logro de resultados de aprendizaje a corto plazo, hay otros elementos tan importantes en la formación como el éxito profesional, la satisfacción, la persistencia, así como la adquisición de habilidades y competencias que les serán útiles a lo largo de la vida.

De igual forma, los factores y las relaciones resultantes de este análisis, aportan elementos que contribuyen a la validación del modelo de habilidades socioemocionales de Farrington et al. (2012). Aunque nuestra red semántica muestra los elementos de una forma más desagregadas, se pueden observar interrelaciones muy parecidas a las presentadas por Farrington et al., y se puede inferir que los factores no cognitivos tienen influencia en el aprendizaje y en el desempeño académico de los estudiantes.

## 8.6 Discusión

En este apartado se comparte una reflexión final de los resultados generados en la etapa cuantitativa y cualitativa. Así mismo, se aportan los elementos centrales para poder determinar si la hipótesis y la premisa o supuestos de la investigación, establecidos al inicio del estudio, se han confirmado o no.

En relación con la generación de los modelos predictivos de riesgo académico, como se ha observado en los resultados de la etapa de minería de datos, se obtuvieron 5 modelos, cada uno de los cuales es considerado como buen clasificador, pues sus valores de AUC ROC los ubican en ese nivel, según (Minguillón et al., 2017). Los modelos incluyen distintos tipos de variables, tanto académicas como no académicas, que, de acuerdo con la revisión de la literatura realizada, aportan un mayor poder predictivo al considerar aspectos no necesariamente relacionados con la escuela.

Otro elemento de innovación que aporta el proyecto, es la implementación de los modelos en un sistema informático, lo cual facilita la recuperación de datos, la predicción, el informe de los resultados o la generación de alertas. Lo anterior, para usuarios poco familiarizados con el uso de modelos predictivos, es de mucha ayuda y facilita su empleo en el contexto educativo. Por el contrario, en la mayoría de los trabajos previos se buscaba solamente llegar a la etapa de obtención de modelos, sin llegar a la implementación práctica.

**Verificación de la hipótesis.** Para verificar la hipótesis de la etapa cuantitativa, particularmente de la generación de modelos predictivos, se siguieron los procedimientos indicados por la minería de datos educativa para la validación de modelos. Los modelos predictivos obtenidos indican relaciones significativas entre el rendimiento escolar de alumnos de primer ingreso en carreras del área de computación de la Facultad de Matemáticas de la UADY, con factores académicos, resultados de exámenes de ingreso (Exani-II), y antecedentes socioeconómicos, así como la preparatoria de origen, la edad y el responsable del alumno. Las pruebas

de validación cruzada para el modelo generado por el algoritmo LMT, el cual obtuvo los mejores resultados, indican una precisión de 75.42% y una calidad de 0.805 según el indicador del área bajo la curva ROC, que lo ubica como un buen clasificador.

Por tanto, se verifica que el rendimiento escolar de alumnos de primer ingreso a carreras del área de computación en la Facultad de Matemáticas de la Universidad Autónoma de Yucatán, está relacionado con antecedentes extraescolares y académicos de los estudiantes, y que al ser analizados sirvieron para generar modelos predictivos eficientes.

En la etapa relacionada con la intervención educativa y el uso del sistema predictivo se ha podido verificar el impacto de las estrategias desarrolladas y sus beneficios en los siguientes aspectos:

- Detección oportuna de estudiantes en riesgo académico, tanto a nivel individual como a nivel de grupo. De acuerdo con la opinión de los académicos y directivos, la información proporcionada por el sistema, referente a la detección de riesgo académico, permite identificar alumnos y grupos que requieren de medidas de intervención; con lo cual se tuvo la oportunidad de tomar medidas preventivas para apoyar a los alumnos, por ejemplo, pláticas con sus tutores, diseño de cursos a la medida, o la sugerencia para tomar otros cursos de regularización.
- Implementación de un procedimiento de seguimiento o monitoreo en el sistema, para conocer la opinión de los alumnos sobre los cursos que están tomando. Con la información generada, los maestros tuvieron elementos para ajustar planes de clase y estrategias docentes; además, para los alumnos, fue un elemento que facilitó sus procesos metacognitivos referentes a su adaptación a la escuela y su desempeño académico.
- Mejora en los niveles de autodirección del grupo intervenido, al pasar de un nivel Bajo a Moderado, el grupo paso de un nivel de autodirección inicial de

39.3%, a uno final de 57.2%. Por lo que se obtuvo un efecto moderado en el incremento del perfil de la autodirección ( $r=0.333$ ).

- Disminución de alumnos con bajas calificaciones (promedio entre 0 y 69), al reducir de un 57%, según los resultados históricos (2016-2018), a un 22%, en el grupo intervenido (2021).
- Disminución de alumnos con materias reprobadas, en el grupo intervenido el 26% de los alumnos tuvo 2 o más materias reprobadas al finalizar el semestre (2021), mientras que en los resultados históricos (2016-2018) el 64% reprobó dos o más asignaturas.
- Los comentarios de los alumnos reflejan una maduración de procesos metacognitivo, emotivos y actitudinales, además de sentirse más seguros en el manejo de las TIC, la comunicación y la colaboración, entre otros aspectos del aprendizaje logrado en el curso.

De acuerdo con los resultados y reflexiones realizadas, se puede afirmar que la información, los modelos y sistemas generados durante el desarrollo del proyecto de investigación, han servido como instrumentos de apoyo para la planificación educativa y la toma de decisiones oportuna de tutores, docentes, directivos y alumnos, lo cual ha impactado de manera positiva en el rendimiento académico de los estudiantes.

El trabajo realizado está en consonancia con otras propuestas (Silva, 2011; Velázquez y Rodríguez, 2014), que destacan la importancia de realizar investigaciones educativas e intervenciones en los primeros años universitarios, para mejorar las expectativas de éxito académico y facilitar la transición a la universidad.

El enfoque holístico del proyecto, el cual considera múltiples elementos para su desarrollo a la vez que emplea una gran variedad de estrategias de mejora, incluyendo a la tecnología educativa, refuerza los planteamientos de Dorio (2016), en el sentido de que en la transición universitaria y el aprendizaje intervienen

múltiples y complejos factores, y para su entendimiento se requiere de una visualización holística. De igual forma, los resultados también abonan a los señalamientos de Masino y Niño (2016), al afirmar que para planificar intervenciones educativas eficientes, que propicien la mejora del rendimiento y el aprendizaje, es necesario considerar dos o más elementos impulsores del cambio, tomando en cuenta los distintos niveles que conforman el entorno del estudiante.

Si bien los resultados son alentadores, aún hay aspectos que pueden ser mejorados, por lo que, de acuerdo con la metodología de la investigación-acción (Ruiz-Bernardo et al., 2018), será necesario llevar a cabo más iteraciones siguiendo las fases de esta metodología. Esta vez, partiendo de los hallazgos y reflexiones resultantes del proyecto, se realizará el diagnóstico, la planificación y finalmente se implementará el plan de intervención con las mejoras propuestas.

## 9. CONCLUSIONES

Durante el desarrollo del proyecto de tesis se abordaron diversos aspectos para dar respuestas a las preguntas de investigación planteadas y alcanzar los objetivos del estudio. El punto de partida del trabajo fue la identificación de una problemática de bajo rendimiento académico en los alumnos de nuevo ingreso a la Facultad de Matemáticas de la UADY, lo que dio pie a la búsqueda de las causas que influyen en sus resultados educativos, y al diseño de estrategias de intervención educativa, para mitigar la problemática detectada. Además, se deseaba aprovechar la información generada por los sistemas de información institucionales utilizando nuevas metodologías y herramientas basadas en la minería de datos educativos. Lo anterior motivó el planteamiento de las siguientes preguntas que dirigieron la investigación:

1. ¿Cómo entender, predecir el rendimiento escolar de los estudiantes de nuevo ingreso de la Facultad de Matemáticas de la UADY, utilizando herramientas de software?
2. ¿Qué factores extraescolares o académicos determinan significativamente el rendimiento académico de los estudiantes de la Facultad de Matemáticas de la UADY?
3. ¿Qué estrategias de intervención educativa son apropiadas para modificar, mejorar el rendimiento escolar de alumnos con antecedentes problemáticos en la Facultad de Matemáticas de la UADY?

Con los resultados obtenidos en las diferentes etapas del proyecto, se han podido dar respuesta a todas las interrogantes. Las preguntas 1 y 2 son explicadas con los hallazgos del proceso de aplicación de la minería de datos educativos, la implementación de los modelos predictivos en un sistema informático y la información obtenida por el estudio diagnóstico. Como producto de estas etapas se generaron modelos descriptivos y predictivos de la población en estudio, que permitieron el desarrollo del sistema mediante el cual se pudieron identificar grupos

y alumnos en posible riesgo escolar, haciendo uso de datos digitales recuperados de los sistemas y bases de datos institucionales con métodos de la minería de datos educativas.

La pregunta 3, también puede ser contestada con los hallazgos del estudio diagnóstico y la etapa de diseño e implementación de las estrategias de intervención educativa. El desarrollo de la intervención fue dirigido por las necesidades detectadas y estuvo focalizada en un grupo que se identificó como de mayor riesgo académico. Los resultados y comentarios de los involucrados indican una mejora de los niveles de rendimiento escolar y de las habilidades de autodirección del grupo intervenido.

Respecto a los objetivos específicos del estudio, se presentan las aportaciones y la descripción de cómo los resultados ratifican su consecución:

OE1. Diseñar modelos predictivos eficientes de rendimiento escolar con base en los antecedentes extraescolares y académicos de los estudiantes, datos recuperados de los sistemas de información institucional.

- La aplicación de la minería de datos educativos permitió el análisis de diversas fuentes de datos del entorno local para extraer conocimiento útil. Los modelos sintetizaron la influencia de diversos factores académicos y extraescolares en el bajo rendimiento académico de los estudiantes, de tal forma que fue posible predecir tal condición de riesgo.
- El uso de nuevas tecnologías, como la minería de datos educativos, aceleró y facilitó el análisis de información histórica de los alumnos y del contexto académico en que se desenvuelven.
- La eficiencia y precisión de los modelos fue determinada mediante métodos de validación cruzada, y de acuerdo con (Minguillón et al., 2017), los modelos generados por las diversas técnicas se identifican como *Buenos*, por tanto, brindan la confianza para ser utilizados en la predicción de nuevos casos.

OE2. Implementar modelos predictivos mediante un sistema de software que facilite el análisis de nuevas instancias de alumnos y genere informes de predicción de riesgo académico.

- Se desarrolló un sistema de software que implementó los modelos y algoritmos predictivos, generados previamente, para analizar grupos de alumnos de nuevo ingreso y determinar si presentan condiciones de riesgo académico.
- El sistema establece una conexión con la plataforma Moodle y facilita la recuperación de datos de encuestas con las variables requeridas para aplicar los modelos predictivos, a su vez genera informes y permiten el envío de mensajes personalizados a los alumnos. Esta simple acción de notificar una condición de riesgo, según Jayaprakash et al. (2014), es una medida de intervención que puede afectar significativamente el comportamiento del estudiante y hacer que busquen ayuda.
- Los informes de riesgo fueron utilizados para dirigir las estrategias de intervención educativa a los grupos y alumnos identificados como prioritarios para ser atendidos.
- Adicionalmente, se implementó un módulo de seguimiento en el sistema, el cual recupera datos de encuestas de satisfacción de los cursos que llevan los alumnos y genera gráficas con las tendencias. Esta herramienta y sus resultados, permite a los alumnos un proceso reflexivo que coadyuva a la metacognición en su proceso educativo; de igual forma, a los maestros les es de utilidad para tener una retroalimentación de su planeación y estrategias pedagógicas, con lo cual pueden emprender acciones de mejora; a los responsables de la gestión educativa les puede ayudar en la identificación de situaciones que requieren de su intervención, lo anterior es consistente con las estrategias y hallazgos de Velázquez y Rodríguez (2014).

OE3. Identificar elementos del contexto académico que puedan afectar la transición de los estudiantes a la universidad y que condicionen su desempeño académico.

- El análisis diagnóstico permitió identificar los elementos y factores que deberían ser atendidos por las medidas de intervención educativa. Específicamente, se vio la necesidad de madurar sus habilidades académicas, atendiendo aspectos cognitivos y de autodirección del aprendizaje y también de fortalecer o ampliar los mecanismos de tutoría académica.
- También, el contar con modelos y sistemas predictivos permitió detectar, con mayor oportunidad, condiciones y alumnos en riesgo para desarrollar procesos de intervención educativa más eficaces.

OE4. Definir estrategias de intervención educativa para los alumnos con probable bajo rendimiento académico, con énfasis en la utilización de tecnología educativa.

- Tomando como referencia las necesidades de los alumnos de nuevo ingreso, se diseñó un plan de intervención cuyas acciones fueron guiadas mediante el desarrollo de un curso impartido en modalidad virtual, el enfoque seguido es similar a la propuesta de Silva (2011) con los seminarios de primer año, o la de Velázquez y Rodríguez (2014), con su concepto de Tutoría de Asignatura.
- La utilización de tecnología educativa fue fundamental para poder impartir el curso, más aún en el contexto de la pandemia debida al Covid-19, que obligó el cierre físico de las instituciones educativas y la transición a la modalidad virtual (*e-learning*).
- El uso de diversas estrategias de aprendizaje mediadas con herramientas TIC, contribuyó a la motivación y el aprendizaje significativo y colaborativo,

además de facilitar el acceso a los recursos desde cualquier ubicación y en cualquier momento.

OE5. Evaluar la efectividad de las estrategias de intervención implementadas en la búsqueda de la mejora continua de los procesos educativos.

- El uso del cuestionario CIPA+ permitió determinar el impacto del plan de intervención en los componentes del perfil autodirigido de los estudiantes.
- En análisis y seguimiento de las calificaciones y las materias aprobadas aportaron datos para contrastar los resultados históricos con los resultados obtenidos por los alumnos que participaron en la intervención.
- En general, la implementación de las estrategias de intervención ha mostrado resultados favorables en el desarrollo de habilidades académicas y en el rendimiento académico.

Con los elementos descritos anteriormente, en un plano global, se puede decir que el objetivo general del proyecto de tesis se ha cumplido, dicho objetivo fue: Desarrollar un sistema predictivo de rendimiento escolar de alumnos de primer ingreso a carreras del área de computación en la Facultad de Matemáticas de la Universidad Autónoma de Yucatán (UADY), utilizando modelos generados mediante minería de datos educativa, para la detección y atención oportuna de problemas asociados al bajo rendimiento académico.

Gracias a los resultados del estudio se tiene un conocimiento más preciso de la población estudiada, pues se logró caracterizar a los estudiantes de acuerdo con diversos factores académicos y no académicos, estos conocimientos y reglas predictivas quedaron sintetizados en los modelos obtenidos. El sistema posibilita la automatización de la identificación de alumnos en riesgo académico, con lo que se tienen mejores oportunidades de corregir o incidir en los procesos de aprendizaje, de una manera más oportuna y eficiente.

Los resultados también pueden ser de utilidad para otras poblaciones con características y contextos similares, la metodología puede ser adaptada para desplegar estrategias de intervención como las desarrolladas en este estudio. La propuesta también ha logrado alcanzar las metas de la intervención, ya que el instrumento utilizado para medir su impacto, han evidenciado el incremento en el nivel de autodirección de los alumnos y se ha mejorado el rendimiento académico.

Como trabajo futuro, se dará continuidad al proceso de análisis del impacto de las intervenciones al considerar un mayor número de alumnos y grupos para implementar las estrategias de intervención, además de considerar calificaciones y elementos cualitativos para complementar los resultados. También se recomienda ampliar el sistema de software con informes y gráficas que permitan el seguimiento longitudinal de los estudiantes y de los cursos, además de evaluar su usabilidad.

## 10. REFERENCIAS

- Aceves, N. (2008). *Adaptación, Confiabilidad y Validez del Cuestionario de Indagación del Perfil Autodirigido (CIPA) y su evaluación en adultos jóvenes que pertenecen a la Sociedad (Tesis Doctoral)*. Instituto Tecnológico y de Estudios Superiores de Monterrey, México.
- Albarracín, Á. P., & Montoya, D. A. (2016). Programas de intervención para Estudiantes Universitarios con bajo rendimiento académico. *Informes Psicológicos*, 16(1), 13–34. <https://doi.org/10.18566/infpsicv16n1a01>
- Aldikanji, E., & Ajami, K. (2016). Studying Academic Indicators within Virtual Learning Environment Using Educational Data Mining. *International Journal of Data Mining & Knowledge Management Process*, 6(6), 29–42. <https://doi.org/10.5121/ijdkp.2016.6603>
- Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, 37, 13–49. <https://doi.org/10.1016/j.tele.2019.01.007>
- Alonso, C., Gallego, D., & Honey, P. (1999). *Los Estilos de Aprendizaje: Procedimientos de diagnóstico y mejora (7a. Ed.)* (Issue December). Ediciones Mensajero.
- Álvarez, A. C., & Álvarez, V. (2014). *Métodos en la Investigación Educativa* (2 ed.). UPN.
- Alyahyan, E., & Düştögör, D. (2020). Predicting academic success in higher education: literature review and best practices. *International Journal of Educational Technology in Higher Education*, 17(1). <https://doi.org/10.1186/s41239-020-0177-7>
- Anoopkumar, M., & Rahman, A. M. J. (2016). A Review on Data Mining techniques and factors used in Educational Data Mining to predict student amelioration. *Proceedings of 2016 International Conference on Data Mining and Advanced Computing, SAPIENCE 2016*, 122–133. <https://doi.org/10.1109/SAPIENCE.2016.7684113>
- Aziz, A. A., Hafieza, N., & Ahmad, I. (2014). First Semester Computer Science Students ' Academic Performances Analysis by Using Data Mining Classification Algorithms. *Proceeding of the International Conference on Artificial Intelligence and Computer Science(AICS 2014), September*, 100–109.
- Baker, R. S., & Inventado, P. S. (2014). Chapter 4 Educational Data Mining and Learning Analytics. In J. A. Larusson & B. White (Eds.), *Learning Analytics: From Research to Practice* (Issue December, pp. 61–75). Springer.
- Baker, R. S., Lindrum, D., Lindrum, M. J., & Perkowski, D. (2015). Analyzing Early At-Risk Factors in Higher Education e-Learning Courses. *Proceedings of the*

- 8th International Conference on Educational Data Mining (EDM 2015)*, 150–155.
- Bakhshinategh, B., Zaiane, O. R., ElAtia, S., & Ipperciel, D. (2018). Educational data mining applications and tasks: A survey of the last 10 years. *Education and Information Technologies*, 23(1), 537–553. <https://doi.org/10.1007/s10639-017-9616-z>
- Ballester, L., Nadal, A., & Amer, J. (2017). *Métodos y técnicas de investigación educativa* (2 ed.). Ediciones UIB.
- Barraza, A. (2010). *Elaboración de Propuestas de Intervención Educativa*. Universidad Pedagógica de Durango.
- Bates, A. W. (2015). *Teaching in a digital age: Guidelines for designing teaching and learning*. Tony Bates Associates Ltd.
- Benítez, M. G., Barajas, J. I., & Hernández, I. N. (2014). Efecto de la aplicación de una estrategia de comprensión de lectura en un entorno virtual. *Revista Electronica de Investigacion Educativa*, 16(3), 71–87.
- Berens, J., Schneider, K., Görtz, S., Oster, S., & Burghoff, J. (2019). Early Detection of Students at Risk-Predicting Student Dropouts Using Administrative Student Data from German Universities and Machine Learning Methods. *Journal of Educational Data Mining*, 11(3), 1–41.
- Bouckaert, R. R., Frank, E., Hall, M., Kirkby, R., Reutemann, P., Seewald, A., & Scuse, D. (2018). *WEKA Manual for Version 3-8-3*. The University of Waikato.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Brooks, D. C. (2016). *ECAR Study of Undergraduate Students and Information Technology, 2016*. <https://doi.org/10.4324/9781315562049>
- Buenaño-Fernández, D., Gil, D., & Luján-Mora, S. (2019). Application of Machine Learning in Predicting Performance for Computer Engineering Students: A Case Study. *Sustainability*, 11(10), 2833. <https://doi.org/10.3390/su11102833>
- Burgos, C., Campanario, M. L., Peña, D., Lara, J., Lizcano, D., & Martínez, M. (2018). Data mining for modeling students' performance: A tutoring action plan to prevent academic dropout. *Computers and Electrical Engineering*, 66, 541–556. <https://doi.org/10.1016/j.compeleceng.2017.03.005>
- Cabero, J. (2003). Replanteando la tecnología educativa. *Comunicar: Revista Científica Iberoamericana de Comunicación y Educación*, 21, 23–30. <https://doi.org/10.3916/25555>
- Cabero, J. (2015). Reflexiones educativas sobre las tecnologías de la información y la comunicación (TIC). *Revista Tecnología, Ciencia y Educación*, 1(1), 19–27.
- Cabero, J. (2016). ¿Qué debemos aprender de las pasadas investigaciones en Tecnología Educativa? *Revista Interuniversitaria de Investigación En Tecnología Educativa*, 0, 23–33. <https://doi.org/10.6018/riite/2016/256741>
- Carpio, C., Pacheco, V., Carpio, C., Morales, G., Carranza, J., Chaparro, M., Rodríguez, R., Canales, C., Chávez, E., Aguilar, F., & García, D. (2018). Riesgo Académico: Un modelo de intervención, evidencias y extensiones. *Avances Conceptuales y Metodológicos En La Atención al Riesgo Académico: Una Experiencia*, 1–23.

- Cázares, Y. (2002). *Hacia un modelo de componentes que explican el aprendizaje autodirigido en estudiantes adultos mexicanos en cursos en línea de la Universidad TecMilenio*. Universidad Virtual del Tecnológico de Monterrey.
- CENEVAL. (2021). *El EGEL Plus y el Proyecto Egresos de Licenciatura*. Centro Nacional de Evaluación para la Educación Superior. [https://ceneval.edu.mx/wp-content/uploads/2021/09/EGEL-PLUS\\_DIGITAL-1.pdf](https://ceneval.edu.mx/wp-content/uploads/2021/09/EGEL-PLUS_DIGITAL-1.pdf)
- Chaves, E., Trujillo, J. M., & López, J. A. (2015). Autorregulación del aprendizaje en entornos personales de aprendizaje en el grado de educación primaria de la universidad de granada, España. *Formacion Universitaria*, 8(4), 63–76. <https://doi.org/10.4067/S0718-50062015000400008>
- Chong, E. (2017). Factores que inciden en el rendimiento académico de los estudiantes de la Universidad Politécnica del Valle de Toluca. *Revista Latinoamericana de Estudios Educativos*, 47(1), 91–108.
- Costa, E. B., Fonseca, B., Santana, M. A., de Araújo, F. F., & Rego, J. (2017). Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Computers in Human Behavior*, 73, 247–256. <https://doi.org/10.1016/J.CHB.2017.01.047>
- Creswell, J. (2012). *Educational research (4th ed.): planning, conducting, and evaluating quantitative and qualitative research*. Pearson.
- Creswell, J. (2014). *Research design (4th ed.): Qualitative, quantitative, and mixed methods approaches*. SAGE.
- Díaz Barriga, F., & Hernández Rojas, G. (2010). *Estrategias docentes para un aprendizaje significativo*. McGraw-Hill.
- Dorio, I. (2016). *La transición a la Universidad. El grado de maestro de Educación Infantil (Tesis Doctoral)*. Universitat de Barcelona, España.
- Elcullada, R., Galang, A. A., & Hallar, B. J. (2021). The Impact and Effectiveness of E-Learning on Teaching and Learning. *International Journal of Computing Sciences Research*, 5(1), 383–397. <https://doi.org/10.25147/ijcsr.2017.001.1.47>
- Escamilla, M. A., & Heredia, Y. (2019). Autodirección, habilidades de pensamiento y rendimiento académico en estudiantes normalistas. *Diálogos Sobre Educación*, 19, 1–21. <https://doi.org/10.32870/dse.v0i19.492>
- Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Seneca Keyes, T., Johnson, D. W., & Beechum, N. O. (2012). *Teaching adolescents to become learners. The role of noncognitive factors in shaping school performance: A critical literature review*. Chicago: University of Chicago Consortium on Chicago School Research.
- Fawcett, T. (2003). ROC Graphs: Notes and Practical Considerations for Data Mining Researchers. *HP Invent*, 27. <https://doi.org/10.1.1.10.9777>
- Fernández-Pampillón, A. (2009). Las plataformas e-learning para la enseñanza y el aprendizaje universitario en Internet. In *Las plataformas de aprendizaje. Del mito a la realidad*. Biblioteca Nueva.

- Fox, A., Patterson, D. A., Ilson, R., Joseph, S., Walcott-Justice, K., & Williams, R. (2014). *Software Engineering Curriculum Technology Transfer: Lessons learned from MOOCs and SPOCs*. UCB/EECS-2014-17.
- García, D. (2015). *Construcción de un Modelo para Determinar el Rendimiento Académico de los Estudiantes Basado en Learning Analytics (Análisis del Aprendizaje), mediante el Uso de Técnicas Multivariantes (Tesis Doctoral)*. Universidad de Sevilla, España.
- Gómez, M., Galeano, C., & Jaramillo, D. A. (2015). El estado del arte: una metodología de investigación. *Revista Colombiana de Ciencias Sociales*, 6(2), 423–442.
- Gros, B. (2012). Retos y tendencias sobre el futuro de la investigación acerca del aprendizaje con tecnologías digitales. *RED. Revista de Educación a Distancia*, 32.
- Guerra, M. D., & Borrillo, A. (2018). Tutoring and academic performance from the perspective of Health Sciences students and teachers. A systematic review. *Educacion Medica*, 19(5), 301–308.  
<https://doi.org/10.1016/j.edumed.2017.03.019>
- Hamoud, A. K., Hashim, A. S., & Awadh, W. A. (2018). Predicting Student Performance in Higher Education Institutions Using Decision Tree Analysis. *International Journal of Interactive Multimedia and Artificial Intelligence*, 5(2), 26. <https://doi.org/10.9781/ijimai.2018.02.004>
- Hernández, R., Fernández, C., & Baptista, P. (2014). *Metodología de la Investigación 6a Ed.* McGraw-Hill.
- Hernández Rojas, G. (1997). *Módulo Fundamentos del Desarrollo de la Tecnología Educativa (Bases Psicopedagógicas)*. Coordinador: Frida Díaz Barriga Arceo.
- Hidalgo, N., & Murillo, F. J. (2017). Las Concepciones sobre el Proceso de Evaluación del Aprendizaje de los Estudiantes. *REICE. Revista Iberoamericana Sobre Calidad, Eficacia y Cambio En Educación*, 15.1(2017), 107–128. <https://doi.org/10.15366/reice2017.15.1.007>
- Imran, M., Latif, S., Mehmood, D., & Shah, M. S. (2019). Student Academic Performance Prediction using Supervised Learning Techniques. *International Journal of Emerging Technologies in Learning (IJET)*, 14(14), 92–104. <https://doi.org/https://doi.org/10.3991/ijet.v14i14.10310>
- INTEF. (2017). *Marco común de competencia digital docente*.
- Jayaprakash, S. M., Moody, E. W., Lauría, E. J. M., Regan, J. R., & Baron, J. D. (2014). Early Alert of Academically At-Risk Students: An Open Source Analytics Initiative. *Journal of Learning Analytics*, 1(1), 6–47. <https://doi.org/10.18608/jla.2014.11.3>
- Ji, H., Park, K., Jo, J., & Lim, H. (2016). Mining students activities from a computer supported collaborative learning system based on peer to peer network. *Peer-to-Peer Networking and Applications*, 9(3), 465–476. <https://doi.org/10.1007/s12083-015-0397-0>
- Kerlinger, F. N., & Lee, H. (2002). *Investigación del comportamiento (4a ed.)*. McGraw-Hill.

- Kumar, M., & Singh, A. J. (2017). Evaluation of Data Mining Techniques for Predicting Student's Performance. *International Journal of Modern Education and Computer Science*, 8, 25–31. <https://doi.org/10.5815/ijmecs.2017.08.04>
- Kumar, M., Singh, A. J., & Handa, D. (2017). Literature Survey on Student's Performance Prediction in Education using Data Mining Techniques. *International Journal of Education and Management Engineering*, 7(6), 40–49. <https://doi.org/10.5815/ijeme.2017.06.05>
- Lamas, H. (2015). Sobre el rendimiento escolar. *Propósitos y Representaciones*, 3(1), 351–386. <https://doi.org/10.20511/pyr2015.v3n1.74>
- Landwehr, N., Hall, M., & Frank, E. (2006). Logistic model trees. *Machine Learning*, 2837, 241–252. [https://doi.org/10.1007/978-3-540-39857-8\\_23](https://doi.org/10.1007/978-3-540-39857-8_23)
- Le Cessie, S., & Van Houwelingen, J. C. (1992). Ridge Estimators in Logistic Regression. *Applied Statistics*, 41(1), 191–201.
- López, C. E., Guzmán, E. L., & González, F. A. (2015). A Model to Predict Low Academic Performance at a Specific Enrollment Using Data Mining. *Revista Iberoamericana de Tecnologías Del Aprendizaje*, 10(3), 119–125. <https://doi.org/10.1109/RITA.2015.2452632>
- López-Ramírez, V. M. (2015). *Método sistémico para evaluar el rendimiento académico en instituciones de educación superior (Tesis Doctoral)*. Instituto Politécnico Nacional, México.
- Machado, M., & Ochoa, L. A. (2017). Análisis del perfil autodirigido en estudiantes adultos que participan en cursos , diplomados y otros programas online de desarrollo profesional, en el contexto de América Latina. *XIV Congreso Nacional de Investigación Educativa*, 1–14.
- Marques, J., Hobbs, D., & Graf, S. (2014). Integrating an at-risk student model into learning management systems. *Nuevas Ideas En Informática Educativa TISE*, 120–124.
- Márquez-Vera, C., Romero, C., & Ventura, S. (2012). Predicción del Fracaso Escolar Mediante Técnicas de Minería de Datos. *IEEE-Rita*, 7(3), 109–117.
- Martínez, D. L., Karanik, M., Giovannini, M., & Pinto, N. (2015). Perfiles de Rendimiento Académico: Un Modelo basado en Minería de datos. *Campus Virtuales*, 6(1), 12–30.
- Masino, S., & Niño, M. (2016). What works to improve the quality of student learning in developing countries? *International Journal of Educational Development*, 48, 53–65. <https://doi.org/10.1016/j.ijedudev.2015.11.012>
- Menacho, C. H. (2017). Predicción del rendimiento académico aplicando técnicas de minería de datos. *Anales Científicos*, 78(1), 26. <https://doi.org/10.21704/ac.v78i1.811>
- Merchan, S. M., & Duarte, J. A. (2016). Analysis of Data Mining Techniques for Constructing a Predictive Model for Academic Performance. *IEEE Latin America Transactions*, 14(6), 2783–2788. <https://doi.org/10.1109/TLA.2016.7555255>
- Miguéis, V. L., Freitas, A., Garcia, P. J. V., & Silva, A. (2018). Early segmentation of students according to their academic performance: A predictive modelling

- approach. *Decision Support Systems*, 115, 36–51.  
<https://doi.org/10.1016/j.dss.2018.09.001>
- Minguillón, J., Casas, J., & Minguillón, J. (2017). *Minería de datos: modelos y algoritmos*. Editorial UOC.
- Miranda, M. A., & Guzmán, J. (2017). Análisis de la deserción de estudiantes universitarios usando técnicas de minería de datos. *Formacion Universitaria*, 10(3), 61–68. <https://doi.org/10.4067/S0718-50062017000300007>
- Mitra, S., & Pal, S. K. (1995). Fuzzy multi-layer perceptron, inferencing and rule generation. *IEEE Transactions on Neural Networks*, 6(1), 51–63.
- Molina, M. (2015). Valoración de los criterios referentes al rendimiento académico y variables que lo puedan afectar. *Revista Médica Electrónica*, 37(6), 617–626.
- Monsalve, H. de J. (2016). *Aproximaciones hacia una definición de “Bajo Rendimiento Escolar.”*
- Montes, I. C., & Lerner, J. (2010). Rendimiento Académico de los estudiantes de pregrado de la Universidad EAFIT. *Perspectiva Cuantitativa*, 158.
- Moreno, J. A., Silveira, Y., & Alias, A. (2015). Modelo predictivo para la mejora de la percepción de competencia y rendimiento académico en estudiantes universitarios. *REDU. Revista de Docencia Universitaria*, 13(2), 173–188. <https://doi.org/10.4995/redu.2015.5443>
- Muñoz, A. (2015). *Modelos para la Mejora del Rendimiento Académico de Alumnos de la E.S.O. mediante Técnicas de Minería de Datos (Tesis Doctoral)*. Universidad de Murcia, España.
- Navarro, E. R. (2003). El rendimiento académico: concepto, investigación y desarrollo. *REICE. Revista Iberoamericana Sobre Calidad, Eficacia y Cambio En Educación*, 1(2), 1–15.
- OECD. (2018). “México.” In *Education at a glance: OECD indicators*. OECD Publishing. <https://doi.org/doi.org/10.1787/kppaex-es>
- Orjuela, A., & Rojas, M. (2008). Las Metodologías de Desarrollo Ágil como una Oportunidad para la Ingeniería del Software Educativo. *Revista Avances En Sistemas e Informática*, 5(2), 159–171.
- Padua, L. M. (2019). Factores individuales y familiares asociados al bajo rendimiento académico en estudiantes universitarios. *Revista Mexicana de Investigación Educativa*, 24(80), 173–195.
- Peña-Ayala, A. (2014). Educational Data Mining. In *Studies in Computational Intelligence* (Vol. 524). <https://doi.org/10.1007/978-3-319-02738-8>
- Pía Otero, M., & Pais, E. (2017). *Las habilidades socioemocionales en el modelo de cinco factores: Un estudio con alumnos participantes de FEPBA 2014*. Unidad de Evaluación Integral de la Calidad y Equidad Educativa, Ciudad de Buenos Aires.
- Pozo, J. I. (2008). La psicología cognitiva del aprendizaje. In *Aprendices y Maestros*. Alianza.
- Quinlan, R. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers.

- Rhode, J., Richter, S., Gowen, P., Miller, T., & Wills, C. (2017). Understanding faculty use of the learning management system. *Online Learning Journal*, 21(3), 68–86. <https://doi.org/10.24059/olj.v%vi%i.1217>
- Rico, A., & Sánchez, D. (2018). Diseño de un modelo para automatizar la predicción del rendimiento académico en estudiantes del IPN / Design of a model to automate the prediction of academic performance in students of IPN. *RIDE Revista Iberoamericana Para La Investigación y El Desarrollo Educativo*, 8(16), 246–266. <https://doi.org/10.23913/ride.v8i16.340>
- Río-Jenaro, C., Calle, R., Martín, E., & Robaina, N. (2018). Rendimiento académico en educación superior y su asociación con la participación activa en la plataforma Moodle. *Estudios Sobre Educacion*, 34, 177–198. <https://doi.org/10.15581/004.34.177-198>
- Ruiz-Bernardo, P., Sánchez-Tarazaga, L., & Mateu-Pérez, R. (2018). La innovación pedagógica de la mano de la investigación-acción para mejorar la calidad de las prácticas externas de los Grados de Maestro/a en Educación Primaria y Educación Infantil. *Revista Electrónica Interuniversitaria de Formación Del Profesorado*, 21(1), 33–49. <https://doi.org/dx.doi.org/10.6018/reifop.21.1.277681>
- Sancho, J. M., Bosco, A., Alonso, C., & Sánchez, J. A. (2015). La política educativa TIC de la Comunidad de Madrid (España): la perspectiva del profesorado. *Revista Latinoamericana de Tecnología Educativa*, 13(2), 1–16. <https://doi.org/10.17398/1695>
- Schunk, D. (2012). *Teorías del aprendizaje. Una perspectiva educativa* (6th ed.). Pearson Educación.
- Shahiri, A. M., Husain, W., & Rashid, N. A. (2015). A Review on Predicting Student's Performance Using Data Mining Techniques. *Procedia Computer Science*, 72, 414–422. <https://doi.org/10.1016/j.procs.2015.12.157>
- Sharples, M., Adams, A., Ferguson, R., Gaved, M., McAndrew, P., Rienties, B., Weller, M., & Whitelock, D. (2014). *Innovating Pedagogy 2014: Open University Innovation Report 3*. Milton Keynes: The Open University.
- Silva, M. (2011). El primer año universitario. Un tramo crítico para el éxito académico. *Perfiles Educativos*, 33(Extra 0), 102–114.
- Slater, S., Joksimović, S., Kovanovic, V., Baker, R. S., & Gasevic, D. (2017). Tools for Educational Data Mining: A Review. *Journal of Educational and Behavioral Statistics*, 42(1), 85–106. <https://doi.org/10.3102/1076998616666808>
- Torras, M. E. (2019). *Fundamentos y práctica del aprendizaje adaptativo* (Issue January). Universidad Internacional de Valencia.
- Torres, P. C., & Cobo, J. K. (2017). Tecnología educativa y su papel en el logro de los fines de la educación. *Educere*, 21(68), 31–40.
- Vásquez, J. (2016). *Modelo predictivo para estimar la deserción de estudiantes en una institución de educación superior (Tesis de Magíster)*. Universidad de Chile, Santiago, Chile.
- Velázquez, F. J., & Rodríguez, H. E. (2014). Diseño e instrumentación de una tutoría de asignatura en el programa de licenciatura en Sistemas de Información Administrativa de la Universidad de Guanajuato. *Revista*

- Iberoamericana de Educación Superior*, 5(14), 41–54.  
[https://doi.org/10.1016/s2007-2872\(14\)70299-9](https://doi.org/10.1016/s2007-2872(14)70299-9)
- Villanueva, A., Moreno, L. G., & Salinas, M. J. (2018). Data mining techniques applied in educational environments: Literature review. *Digital Education Review*, 33, 235–266. <https://doi.org/10.1344/der.2018.33.235-266>
- Williams, P., Schrum, L., Sangrà, A., & Guàrdia, L. (2019). *Fundamentos del diseño técnico-pedagógico en e-learning. Modelos de diseño instruccional*. Universitat Oberta de Catalunya.
- Witten, I., Frank, E., & Hall, M. (2011). *Data Mining: Practical Machine Learning Tools and Techniques (3rd ed.)*. Morgan Kaufmann.
- York, T., Gibson, C., & Rankin, S. (2015). Defining and measuring academic success. *Practical Assessment, Research and Evaluation*, 20(5), 1–20.
- Zeng, W., Ju, S., & Hord, C. (2018). A Literature Review of Academic Interventions for College Students With Learning Disabilities. *Learning Disability Quarterly*, 41(3), 159–169. <https://doi.org/10.1177/0731948718760999>

# 11. ANEXOS

## Anexo 1. Instrumento Diagnóstico

### Factores y estrategias para la mejora del Rendimiento Académico

El siguiente cuestionario tiene el propósito de conocer tu opinión acerca de los factores y las diferentes estrategias institucionales, docentes y personales que pudieran incidir en tu rendimiento académico, particularmente al inicio de tus estudios en la Facultad de Matemáticas. La información es recabada de forma anónima y se utilizará exclusivamente para realizar investigación educativa y para la mejora continua de los programas de estudio de la Facultad.

**Instrucciones:** Lee con atención cada una de las preguntas y contesta conforme a tu opinión.

1. Selecciona la carrera en la que estás inscrito(a) \_\_\_\_\_

- a) Licenciatura en Actuaría
- b) Licenciatura en Matemáticas
- c) Licenciatura en Enseñanza de las Matemáticas
- d) Licenciatura en Ciencias de la Computación
- e) Licenciatura en Ingeniería en Computación
- f) Licenciatura en Ingeniería de Software

### Expectativas

2. El promedio de calificación que espero obtener al finalizar el semestre esta entre:

- a) 90 y 100
- b) 80 y 89
- c) 70 y 79
- d) 0 y 69

3. La cantidad de materias que podría reprobado al finalizar el semestre es de:

- a) Ninguna
- b) 1
- c) 2
- d) 3
- e) 4 o más

### Antecedentes personales

4. Considero que al ingresar a la Facultad cumplía con las condiciones mínimas en cuanto a conocimientos, actitudes y habilidades necesarias para estudiar mi carrera.

- Totalmente de acuerdo
- De acuerdo
- Indeciso
- En desacuerdo
- Totalmente en desacuerdo

5. Siento que el nivel de conocimientos y exigencias de los cursos es más alto del que yo esperaba.

- Totalmente de acuerdo
- De acuerdo
- Indeciso
- En desacuerdo
- Totalmente en desacuerdo

6. Considero que tengo buenos hábitos de estudio, lo que me permite estudiar de manera autónoma y responsable.

- Totalmente de acuerdo
- De acuerdo
- Indeciso
- En desacuerdo
- Totalmente en desacuerdo

7. Tengo situaciones personales, familiares o económicas que afectan mi rendimiento académico.

- Totalmente de acuerdo
- De acuerdo
- Indeciso
- En desacuerdo
- Totalmente en desacuerdo

8. Pienso que necesito más apoyo por parte de la escuela para mejorar mi desempeño académico.

Totalmente de acuerdo  De acuerdo  Indeciso  En desacuerdo  Totalmente en desacuerdo

9. En caso de requerirlo, ¿qué tipo de apoyos necesitas? \_\_\_\_\_

### **Estrategias docentes**

10. Los profesores que imparten clases en el primer semestre de mi carrera tienen la preparación adecuada para impartir bien sus asignaturas.

Totalmente de acuerdo  De acuerdo  Indeciso  En desacuerdo  Totalmente en desacuerdo

11. En los cursos que he llevado se utilizan recursos didácticos y materiales de enseñanza de calidad y adecuados para que aprenda bien la materia.

Totalmente de acuerdo  De acuerdo  Indeciso  En desacuerdo  Totalmente en desacuerdo

12. En la impartición de los cursos que he llevado se emplean variedad de métodos de enseñanza, diferentes al de exposición oral del profesor, (apoyos audiovisuales, multimedios, aulas interactivas, desarrollo de proyectos, prácticas de laboratorio, etc.)

Muy frecuentemente  Frecuentemente  Ocasionalmente  Raramente  Nunca

13. En mis materias se aplican combinaciones de varios mecanismos de evaluación, tales como: exámenes, tareas, problemas para resolver, prácticas de laboratorio, trabajos e informes, etc.

Muy frecuentemente  Frecuentemente  Ocasionalmente  Raramente  Nunca

14. En los cursos que he llevado se emplea la evaluación formativa (evaluación utilizada para retroalimentar el proceso educativo que normalmente no cuenta para la calificación).

Muy frecuentemente  Frecuentemente  Ocasionalmente  Raramente  Nunca

### **Estrategias docentes con apoyo de tecnología educativa**

15. En las materias que he llevado se utiliza la plataforma virtual (cursos en línea) como apoyo importante para el desarrollo de actividades de enseñanza-aprendizaje.

Muy frecuentemente  Frecuentemente  Ocasionalmente  Raramente  Nunca

16. En las materias que utilizan la plataforma virtual (cursos en línea), se emplea varias de sus herramientas para el desarrollo de actividades de enseñanza-aprendizaje y evaluación.

Muy frecuentemente  Frecuentemente  Ocasionalmente  Raramente  Nunca

17. Los profesores utilizan tecnología educativa básica en sus clases (videos, Internet, repositorios de materiales digitales, etc.)

Muy frecuentemente  Frecuentemente  Ocasionalmente  Raramente  Nunca

18. Los profesores utilizan tecnología educativa en sus clases que consideras novedosa o avanzada (clases invertidas, aprendizaje móvil, videos interactivos, realidad virtual/aumentada, juegos educativos, herramientas colaborativas, etc.)

Muy frecuentemente  Frecuentemente  Ocasionalmente  Raramente  Nunca

### Entorno institucional

19. La escuela cuenta con suficientes recursos físicos (biblioteca, aulas, laboratorios, centros de cómputo, etc.), para que pueda desarrollar mis actividades de enseñanza-aprendizaje de la mejor manera.

Totalmente de acuerdo  De acuerdo  Indeciso  En desacuerdo  Totalmente en desacuerdo

20. ¿Qué recursos físicos faltan y debería tener la escuela? \_\_\_\_\_

\_\_\_\_\_

21. El plan de estudios de mi carrera está bien diseñado.

Totalmente de acuerdo  De acuerdo  Indeciso  En desacuerdo  Totalmente en desacuerdo

22. ¿Qué cosas deberían mejorarse en el plan de estudios? \_\_\_\_\_

\_\_\_\_\_

23. El acceso a una beca o apoyo económico desde el primer semestre de estudios, me ayudaría a mejorar mi rendimiento académico.

Totalmente de acuerdo  De acuerdo  Indeciso  En desacuerdo  Totalmente en desacuerdo

24. La impartición de talleres adicionales de apoyo, para reforzar temas de materias complicadas o en las que no tengo buenas bases, me ayudaría a mejorar mi rendimiento académico.

Totalmente de acuerdo  De acuerdo  Indeciso  En desacuerdo  Totalmente en desacuerdo

### Tutorías y asesorías

25. Las sesiones de **tutoría** (dudas administrativas o académicas, apoyo carga de materias, procesos administrativos, etc., con un tutor asignado) me han ayudado para mejorar mi desempeño académico.

Totalmente de acuerdo  De acuerdo  Indeciso  En desacuerdo  Totalmente en desacuerdo

26. Las veces que he acudido a sesiones de **tutoría** con mi tutor(a) asignado son:

a) ninguna vez                      b) 1-2 veces                      c) 3-4 veces                      d) 5 o más veces

27. Las sesiones de **asesoría** (tiempo dedicado de los profesores para aclarar dudas de su materia fuera del horario de clases) me han ayudado para mejorar mi desempeño académico.

Totalmente de acuerdo  De acuerdo  Indeciso  En desacuerdo  Totalmente en desacuerdo

28. Las veces que he acudido a sesiones de **asesoría** con algún profesor(a) para aclarar dudas de su materia son:

a) ninguna vez                      b) 1-2 veces                      c) 3-4 veces                      d) 5 o más veces

29. Comentarios adicionales o sugerencias para lograr una mejora en tu rendimiento académico (a nivel personal, institucional o de estrategias docentes):

\_\_\_\_\_

## Anexo 2. Planeación Didáctica del Curso

DATOS GENERALES DE IDENTIFICACIÓN					
Nombre de la asignatura	Habilidades académicas para el primer año universitario				
Tipo	Libre				
Modalidad	Mixta				
Ubicación	Primer semestre				
Duración total en horas	96	Horas presenciales	32	Horas no presenciales	64
Créditos	6				
Requisitos académicos previos	Ninguno				

### COMPETENCIA DE LA ASIGNATURA

Desarrolla habilidades académicas, mediante el empleo de herramientas tecnológicas, estrategias y técnicas de aprendizaje, para la autorregulación del aprendizaje, la madurez de los procesos cognitivos, la comunicación y la adaptación a su entorno universitario.

<b>Unidad I</b>	El entorno universitario
<b>Competencia</b>	Discutir la forma en que diversos elementos y factores del entorno educativo influyen en el proceso de aprendizaje, de manera crítica y proactiva.

Secuencia de contenidos	Resultados de aprendizaje	Desagrega do de contenidos	Estrategias de enseñanza y aprendizaje	Actividades de enseñanza-aprendizaje			
				Descripción	Recursos TIC	Duración	
						H P	HN P
1. Ambientes virtuales de aprendizaje	Identifica las capacidades de los entornos de aprendizaje para soportar los procesos de aprendizaje mediados por TIC.	1.1 Definición y propósito de los entornos virtuales de aprendizaje (EVA) 1.2 Tipos de EVA y sus características 1.3 Uso de EVA disponibles en la institución.	Aprendizaje guiado  Aprendizaje interactivo  Estudio individual	Profesor: - Exposición - Descripción de actividades  Alumnos: - De manera individual, participa en el foro de discusión "Presentación" y explora las herramientas disponibles en el entorno. Comentar cuales herramientas disponibles en el AVA son sus favoritas y por qué.  - Lectura comprensiva.  - Contestar encuesta sobre estilos de aprendizaje.	- Video-conferencia - Presentación PPT - Curso en línea  - Foro de discusión  - Lección multimedia en el AVA - Cuestionarios de repaso en el AVA - Cuestionario en línea.	2	4

2. Medios de comunicación y colaboración	Reconoce las características y usos de los medios de comunicación como herramientas de aprendizaje y colaboración.	2.1 Basados en texto 2.2 Basados en voz 2.3 Basados en video 2.4 Mixtos	Aprendizaje guiado  Aprendizaje colaborativo  Investigación documental  Estudio individual	Profesor: - Exposición - Elaboración de preguntas - Descripción de actividades  Alumnos: - En equipos de 3 o 4 integrantes investigar y elaborar un organizador gráfico ( <i>cuadro comparativo</i> ), en donde se contrasten las principales características de diversos medios o herramientas de comunicación y colaboración basados en el Internet.  - Lectura comprensiva.	- Videoconferencia - Presentación PPT - Curso en línea - Video sobre competencias digitales - Infografía  - Documento de texto compartido en Notas de Teams  - Lección multimedia en el AVA - Cuestionarios de repaso en el AVA	2	4
3. Los recursos en la escuela	Identifica los recursos institucionales que sirven de apoyo en el proceso educativo desarrollado en la Facultad de Matemáticas.	3.1 Recursos físicos y tecnológicos 3.2 Planta docente 3.3 Recursos bibliográficos 3.4 Apoyos para el estudiante	Aprendizaje guiado  Estudio individual	Profesor: - Exposición - Elaboración de preguntas - Descripción de actividades - Resolución de dudas  Alumnos: Contestar cuestionario sobre recursos de la escuela.  Contestar encuesta periódica del desarrollo de sus materias de primer semestre y de su desempeño como alumno.	- Videoconferencia - Presentación PPT - Curso en línea - Páginas Web de recursos y servicios en la escuela  - Cuestionario ludificado  - Cuestionario en línea	2	4
4. Modalidades de enseñanza	Discute de forma crítica y proactiva, como las diversas modalidades de enseñanza demandan distintas habilidades y requisitos para el logro del aprendizaje.	4.1 Modalidad presencial 4.2 Modalidad no presencial 4.3 Modalidad mixta	Aprendizaje guiado  Investigación documental  Aprendizaje colaborativo	Profesor: - Exposición - Elaboración de preguntas - Descripción de actividades - Resolución de dudas  Alumnos: De forma individual, elabora un ensayo sobre las características principales de las modalidades de educación (incluyendo que recursos se requieren y lo que demanda de los alumnos). Se complementa la actividad compartiendo un cuadro sinóptico del tema a un foro de discusión y comentando dos aportaciones de sus compañeros.	- Videoconferencia - Curso en línea  Documento de texto  Foro de discusión	2	4

<b>Unidad II</b>	Motivación como clave para el éxito
<b>Competencia</b>	Identificar los factores motivacionales y emocionales que influyen en el rendimiento académico universitario, para el logro de las metas de aprendizaje.

Secuencia de contenidos	Resultados de aprendizaje	Desagregado de contenidos	Estrategias de enseñanza y aprendizaje	Actividades de enseñanza-aprendizaje			
				Descripción	Recursos TIC	Duración	
						HP	HNP
1. Factores motivacionales que afectan el aprendizaje	Analiza los factores motivacionales que influyen en el rendimiento académico universitario, considerando las metas de aprendizaje.	1.1 Motivación intrínseca y extrínseca 1.2 Motivación del logro y autoeficacia 1.3 Metas de aprendizaje	-Diseño de actividades de Autoestudio  -Estudio individual -Investigación documental -Uso de organizadores gráficos. -Aprendizaje colaborativo	Profesor: - Elaboración de descripción de secuencia de actividades y materiales  Alumnos: De manera individual y luego colaborativa, desarrollar un estudio de caso en el cual se analicen los factores motivacionales implicados para el logro de una meta de aprendizaje.	- Curso en línea - Infografías - Documento interactivo - Videos  - Lección multimedia en el AVA -Búsqueda en Internet - Mapa mental sobre factores motivacionales - Foro de discusión	4	8
2. Factores emocionales que afectan el aprendizaje	Reconoce los factores emocionales que influyen en el rendimiento académico universitario, considerando las características de la inteligencia emocional.	2.1 Clasificación de las emociones 2.2 Inteligencia emocional 2.3 Habilidades sociales: empatía, persistencia y optimismo	-Diseño de actividades de Autoestudio  -Estudio individual -Investigación documental -Uso de organizadores gráficos. -Aprendizaje colaborativo	Profesor: - Elaboración de descripción de secuencia de actividades y materiales  De manera individual y luego colaborativa, realizar y compartir memes en el cual se integren los factores emocionales que influyen en el rendimiento académico.  Contestar encuesta periódica del desarrollo de sus materias de primer semestre y de su desempeño como alumno.	- Curso en línea - Infografías - Documento interactivo - Videos  - Lección multimedia en el AVA -Búsqueda en Internet - Memes sobre situaciones emocionales - Foro de discusión  - Cuestionario en línea	4	8

<b>Unidad III</b>	Autodirección en el aprendizaje
<b>Competencia</b>	Reconoce estrategias y procesos eficaces de estudio, acordes a su contexto personal y con el apoyo de herramientas de uso práctico.

Secuencia de contenidos	Resultados de aprendizaje	Desagregado de contenidos	Estrategias de enseñanza y aprendizaje	Actividades de aprendizaje			
				Descripción	Recursos TIC	Duración	
						HP	HN P
1. El aprendizaje y sus elementos	Identifica las distintas maneras que un individuo puede aprender en ambientes universitarios, para mejorar las estrategias de estudio.	1.1 Definición del aprendizaje 1.2 Fases del aprendizaje (atención, percepción, adquisición, retención y transferencia) 1.3 Estilos de aprendizaje (Honey y Mumford)	-Diseño de actividades de Autoestudio  -Aprendizaje autónomo y reflexivo -Aprendizaje invertido -Investigación documental -Aprendizaje colaborativo -Aprendizaje invertido	Profesor: - Elaboración de descripción de secuencia de actividades y materiales  Alumnos: De manera individual y después colaborativa elabora una tabla comparativa con las características correspondientes a los diferentes estilos de aprendizaje.	- Infografías - Páginas interactivas - Videos  - Lección multimedia en el AVA -Búsqueda en Internet - Documento de texto: cuadro comparativo sobre estilos de aprendizaje - Espacio de colaboración en Notas MS.	3	6
2. Tipos de aprendizaje	Diferencia las estrategias de estudio, de manera personal, con base en los tipos de aprendizaje que se generan en los procesos de formación universitaria.	2.1 Aprendizaje individual 2.2 Aprendizaje cooperativo 2.3 Aprendizaje colaborativo 2.4 Aprendizaje por descubrimiento (activo)	-Diseño de actividades de Autoestudio  -Aprendizaje autónomo y reflexivo -Investigación documental -Aprendizaje colaborativo -Aprendizaje invertido	Profesor: - Elaboración de descripción de secuencia de actividades y materiales  Alumnos: De manera individual y posteriormente de manera colaborativa, elaborar un video donde se expliquen los tipos de aprendizajes y sus estrategias estudio.  Contestar encuesta periódica del desarrollo de sus materias de primer semestre y de su desempeño como alumno.	- Curso en línea - Infografías - Documento interactivo - Videos  - Lección multimedia en el AVA -Búsqueda en Internet - Video sobre tipos de aprendizaje. - Espacio de colaboración en Notas compartidas	3	6
3. Hábitos de estudios para el aprendizaje	Describe las técnicas y rutinas de estudio en ambientes universitarios, para mejorar el rendimiento académico.	3.1 Características del hábito: deseo, saber y práctica 3.2 Ciclo del hábito: recordatorio, rutina y recompensa 3.3 Hábitos que potencian el aprendizaje:	-Diseño de actividades de Autoestudio  -Aprendizaje autónomo y reflexivo -Investigación documental -Aprendizaje colaborativo -Aprendizaje invertido	Profesor: - Elaboración de descripción de secuencia de actividades y materiales  Alumnos: De manera colaborativa, elaborar un mapa mental donde se plasmen las técnicas y hábitos de estudio para mejorar su rendimiento académico.	- Curso en línea - Infografías - Documento interactivo - Videos  - Lección multimedia en el AVA -Búsqueda en Internet - Video sobre tipos de aprendizaje. - Espacio de colaboración en Notas compartidas	3	6

		Organizar, planear y hacer					
		3.4 Prácticas que afectan los hábitos de estudio: procrastinación académica, desorganización y falta de sueño					

<b>Unidad IV</b>	Mi visión como profesionista
<b>Competencia</b>	Establece metas de desarrollo educativo, a partir del análisis de su carrera y del ámbito profesional, tomando en cuenta variados puntos de vista, conceptos, teorías y explicaciones.

Secuencia de contenidos	Resultados de aprendizaje	Desagregado de contenidos	Estrategias de enseñanza y aprendizaje	Actividades de aprendizaje			
				Descripción	Recursos TIC	Duración	
						HP	HN P
1. Las competencias de egreso	Reflexiona sobre las competencias profesionales de un licenciado en ciencias de la computación, de cara a su futuro laboral.	1.1 Las competencias de egreso de un LCC  1.2 La importancia de las competencias genéricas	-Diseño de actividades de Autoestudio -Aprendizaje guiado  -Aprendizaje autónomo y reflexivo -Investigación documental -Aprendizaje colaborativo	Profesor: - Elaboración de descripción de secuencia de actividades y materiales - Exposición - Elaboración de preguntas  De manera individual y después colaborativa, elaborar un documento en donde se analicen las competencias de egreso de un LCC.	- Curso en línea - Página Web de la carrera - Video-conferencia  - Lección multimedia en el AVA -Búsqueda en Internet - Documento de texto compartido. - Chat - Espacio colaborativo de	3	6
2. Áreas de estudio y especialización	Reflexiona sobre las áreas de desarrollo académico y profesional para enfocar adecuadamente sus motivaciones y esfuerzos.	2.1 Disciplinas de las Ciencias de la Computación  2.2 Ciencias computacionales aplicadas  2.3 Los profesionistas opinan	-Diseño de actividades de Autoestudio -Aprendizaje guiado  -Aprendizaje autónomo y reflexivo -Investigación documental	Profesor: - Elaboración de descripción de secuencia de actividades y materiales - Exposición - Elaboración de preguntas  De manera individual, elaborar un documento en donde se sinteticen las principales ramas o disciplinas de las ciencias de la	- Curso en línea - Videos con entrevistas a profesores y egresados - Video-conferencia  - Lección multimedia en el AVA -Búsqueda en Internet - Documento de texto.	4	8

				<p>computación, también se deberá reflexionar sobre áreas de interés personal en las que quisiera especializarse el alumno.</p> <p>- Elaborar en equipo el portafolio de evidencias final.</p>	<p>- Presentación PPT</p>		
--	--	--	--	--	---------------------------	--	--

## EVALUACIÓN DEL DESEMPEÑO

### EVALUACIÓN DE PROCESO

Estrategia de evaluación	Criterios de evaluación	Ponderación
Organizador gráfico	<ul style="list-style-type: none"> <li>Identifica aspectos centrales de los temas tratados e ideas secundarias</li> <li>Es claro y ordenado, y es fácil de comprender</li> <li>No hay errores gramaticales, ortográficos o de puntuación</li> </ul>	15%
Debates	<ul style="list-style-type: none"> <li>Participa de manera respetuosa y ética, acorde a las reglas establecidas</li> <li>Las aportaciones enriquecen la discusión y el entendimiento del tema</li> <li>Sus aportaciones son claras y acordes a los temas tratados</li> <li>Construye su propia reflexión</li> </ul>	15%
Ensayos	<ul style="list-style-type: none"> <li>Argumenta con claridad y precisión las ideas que propone</li> <li>Organiza la información de manera lógica y coherente</li> <li>Emplea fuentes de información pertinentes y actualizadas</li> <li>Sus conclusiones reflejan asimilación y entendimiento del tema</li> </ul>	15%
Cuestionarios y encuestas	<ul style="list-style-type: none"> <li>Completa la totalidad de los instrumentos aplicados</li> <li>Participa de forma puntual y respetuosa.</li> </ul>	15%

### EVALUACIÓN DE PRODUCTO

Estrategia de evaluación	Criterios de evaluación	Ponderación
Portafolio de evidencias	<ul style="list-style-type: none"> <li>Escribe considerando aspectos ortográficos, gramaticales y de organización.</li> <li>Muestra constancia y esfuerzo, reflejada en la regularidad, cantidad y extensión de los registros realizados.</li> </ul>	40%

- |  |   |  |
|--|---|--|
|  | <ul style="list-style-type: none"><li>• Describe con claridad y objetividad sus ideas</li><li>• Identifica y destaca los elementos más relevantes de los temas analizados</li><li>• Reporta reflexiones valiosas sobre las tareas realizadas, su organización y aspectos emotivos involucrados (esfuerzo, orgullo, satisfacción, sentido de equipo, etc.)</li></ul> |  |
|--|---|--|

## Anexo 3. Características de los Recursos TIC Empleados

### Criterios del uso de recursos TIC y aspectos del aprendizaje que favorecen

Categoría de recurso TIC	Tecnología / Compañía	Intencionalidad educativa / Justificación de uso
<b>Curso en línea</b>	Moodle	Curso implementado en la plataforma de gestión de contenidos educativos Moodle. Integra los recursos didácticos y facilita el acceso a ellos de manera ubicua. Contiene una amplia variedad de herramientas de uso libre, para planificar actividades de aprendizaje y administrar un curso virtual.
<b>Avisos</b>	Moodle - Herramienta Foro de Avisos	Espacio para que el profesor publique noticias relevantes para los participantes del curso. Por ejemplo, avisos de actividades por realizar, sugerencias o apoyos pedagógicos o información sobre cuestiones administrativas que afecten al curso. Además, los mensajes se envían automáticamente a los buzones personales de los alumnos, con lo que se cuenta con un sistema de comunicación eficiente.
<b>Calendario</b>	Moodle - Herramienta Calendario	Es una herramienta para mostrar de forma organizada las actividades y tareas programadas en el curso. El contar con una vista unificada y ordenada por fechas de las tareas pendientes, ayuda a los estudiantes a planificar y enfocar mejor sus esfuerzos.
<b>Lección multimedia en el AVA</b>	Moodle - Herramienta Lección	La actividad Lección presenta una serie de páginas Web con contenidos multimedia (texto, audio, video, gráficas) y elementos para navegar de manera condicional e interactuar.
<b>Cuestionario de repaso en el AVA</b>	Moodle - herramienta de cuestionarios en lecciones	Las páginas tipo Cuestionario de la actividad Lección, presentan preguntas para verificar la comprensión del tema (evaluación formativa), también permite la repetición y navegación condicional.
<b>Presentación PPT</b>	Microsoft PowerPoint	Permite presentar diversos tipos de contenidos para organizar y compartir la información. Facilita la comunicación y el desarrollo del pensamiento.
<b>Documento de texto</b>	Microsoft Word	Facilita la redacción de texto y la inclusión de gráficas, tablas o dibujos, entre otros. Permite la organización y elaboración de ideas, y la compartición de trabajos con un formato claro y legible.  <b>Individual:</b> elaboración individual de trabajos para desarrollar habilidades de escritura, síntesis, organización de ideas, etc.  <b>Compartido:</b> elaboración de forma colaborativa, para desarrollar habilidades de comunicación, trabajo en equipo, responsabilidad, tolerancias, entre otras.
<b>Documento interactivo</b>	Genially	Documento digital que facilita la creación de contenidos interactivos y enriquecidos que posibilitan la navegación no secuencial y con elementos multimedia estéticos. Brinda amplias posibilidades dinámicas que motivan el aprendizaje del estudiante.
<b>Página Web</b>	Diversos enlaces externos	Páginas HTML externas que presentan diversos contenidos relacionados con las temáticas vista en las clases o cursos.

<b>Categoría de recurso TIC</b>	<b>Tecnología / Compañía</b>	<b>Intencionalidad educativa / Justificación de uso</b>
<b>Búsqueda en Internet</b>	Chrome / Firefox / Safari	Facilitan el acceso a diversos recursos disponibles en Internet. Ideal para trabajos que requieran algún tipo de investigación.
<b>Videoconferencia (comunicación síncrona)</b>	Microsoft Teams	Permite crear sesiones de videoconferencia en tiempo real, con diversas opciones para compartir presentaciones, pantallas y comunicarse con grupos o equipos de trabajo. Facilita la comunicación y la colaboración.  * Las sesiones se graban y se colocan en una sección del curso en línea para su consulta de forma asíncrona.
<b>Video (comunicación asíncrona)</b>	Formato jpg. (Filmora / MovieMaker / Loom / Flipgrid )	Es un medio para presentar contenidos de forma ágil y atraer la atención de los estudiantes.  Profesor: puede diseñar videos para exponer temas específicos o recuperar videos de plataformas como Youtube o Vimeo, para reutilizar contenidos generados por otros docentes.  Alumnos: como un medio para expresar sus ideas sobre algún tema, colaborar y presentar resultados de forma ágil y dinámica.
<b>Foro de discusión</b>	Moodle – Herramienta foro	Espacio para presentar ideas, opiniones y debatir en modo asincrónico.
<b>Mapa mental</b>	Draw.io / SimpleMind / CmapTools	Es un tipo de organizador gráfico que representa conceptos relacionados a partir de un tema principal. Su aplicación permite generar, organizar, expresar los aprendizajes y asociar más fácilmente nuestras ideas. Se puede utilizar para evaluar cómo se han procesado los conocimientos.
<b>Infografía</b>	Genially	Interrelación de ideas en redes multimediales de conceptos. Permite resumir un tema para entenderlo con facilidad. Sus elementos visuales atractivos captan la atención y organizan de forma clara la información.
<b>Meme</b>	Formato .jpg - Editor de imágenes o gráficas	Elemento gráfico o fotografía que se modifica para presentar temas, mensajes o expresiones culturales de forma humorística. Ludifica procesos de aprendizaje para captar la atención y motivar a los estudiantes, los sitúa en un contexto cultural determinado.
<b>Espacio de colaboración</b>	Microsoft Teams - Herramienta Notas	Es un espacio al que pueden acceder los miembros de un equipo o un curso para crear o editar o compartir documentos. Facilita el trabajo colaborativo y la comunicación.
<b>Cuestionario / Encuesta en línea</b>	Google forms	Es una herramienta que permite elaborar y aplicar cuestionarios o encuestas, para recopilar información de forma fácil y eficiente, a través de Internet.
<b>Cuestionario ludificado</b>	Quizizz	Cuestionarios presentados con un formato lúdico, se emplean elementos gráficos atractivos y dispone de varias modalidades de competencia, entre otros aspectos, que ayudan a captar la atención de los alumnos y motivarlos.
<b>Chat</b>	Microsoft Teams – Herramienta de chat	Herramienta de comunicación para expresar ideas u opiniones de forma sincrónica o en tiempo real.

## **Anexo 4. Códigos y Transcripciones del Análisis Cualitativo**

### **PI 1: ¿Qué tipo de actividades fueron más significativas para el aprendizaje?**

#### **Áreas de Interés**

Los segmentos de datos asociados a este código son:

1:3 p 5 en E1.pdf: “En esta actividad descubrimos la amplia rama de áreas de interés que nuestra carrera nos ofrece, así como la versatilidad de la misma.”

3:4 p 5 en E3.pdf: “Aprendimos a distinguir las diferentes áreas en las que se divide la carrera de Ciencias de la Computación, así como conocer cuáles son las competencias de egreso que pide la industria, y también tuvimos un acercamiento a las diferentes áreas de especialización que ofrece la carrera.”

3:5 p 5 en E3.pdf: “sirvió como motivante para buscar un mejor desempeño dentro del estudio de las materias, pues conociendo de qué tratan y cuál es el objetivo de esta, permite entender la razón de su estudio y no perder de vista el objetivo.”

4:1 p 3 en E4.pdf: “Esta actividad nos ayudó a mejorar nuestro desempeño en la licenciatura; al motivarnos y mostrarnos lo que podemos conseguir si nos esforzamos. Nos aclaró mucho el panorama de nuestra carrera; ya que como equipo antes de realizar esta actividad realmente no teníamos claras las diversas áreas con las que cuenta nuestra carrera; lo que nos restaba motivación.”

6:3 p 4 en E6.pdf: “podimos tener una idea más acertada de las utilidades que ofrece estudiar nuestra licenciatura, así como las oportunidades de trabajo que nos benefician a largo plazo.”

9:3 p 4 en E9.pdf: “De igual forma, nos llamo la atención los videos en donde nos describen el ámbito laboral en el que podemos desenvolvemos.”

9:4 p 4 en E9.pdf: “El conocer de que se trata la realidad de nuestro ámbito profesional para desarrollar el pensamiento crítico para el análisis de conceptos,

teorías, explicaciones y puntos de vista para fomentar una idea de donde nos desenvolvemos mejor y crear una meta para prepararnos para ese camino.”

### **Competencias de Egreso**

Los segmentos de datos asociados a este código son:

1:1 p 4 en E1.pdf: “nos deja claro las competencias como, que podemos ser, conocer y hacer después de concluir la carrera, como tener un conocimiento general de las distintas áreas de estudio así como prácticas y teóricas.”

2:2 p 3 en E2.pdf: “Nos ayudó a orientarnos en una dirección algo más específica, más que una general, si bien sabíamos a donde queríamos llegar, no sabíamos como, ahora tenemos un camino algo claro de cómo avanzar en las asignaturas, y que es lo importante durante nuestra vida como estudiantes.”

5:3 p 5 en E5.pdf: “aprendimos cuales son todas las competencias que debemos desarrollar en la carrera y cuáles de estas son las que más nos beneficiarán al momento de buscar un trabajo con un buen sueldo.”

7:3 p 4 en E7.pdf: “nos pareció una de las actividades más útiles en cuanto a información, nos brinda una guía para conocer mejor lo que vamos a poder hacer durante y después de terminar nuestra licenciatura. “

7:4 p 4 en E7.pdf: “nos dió un poco más de motivación, viendo lo que podemos hacer como licenciados en ciencias de la computación. Además de empezar a pensar en qué dirección queremos enfocar nuestra carrera.”

8:4 p 5 en E8.pdf: “nos hizo reflexionar si realmente estamos cumpliendo las competencias de la carrera.”

### **Factores Emocionales**

Los segmentos de datos asociados a este código son presentados a continuación:

2:5 p 4 en E2.pdf: “aprendimos a identificar los factores que alteran e integran nuestra vida universitaria ya sea una buena motivación o mala, lo mismo con las emociones, aprendimos a modificarlas y a tratar de controlarlas con la inteligencia emocional.”

3:1 p 3 en E3.pdf: “Aprendimos a identificar los factores motivacionales y emocionales para lograr nuestras metas de aprendizaje.”

4:5 p 5 en E4.pdf: “Los principales aprendizajes que logramos obtener fueron el como podemos canalizar lo que sentimos, el ser responsables emocionalmente y el poder entender que no debemos reprimir nuestras emociones ante cualquier comentario ofensivo.”

5:1 p 3 en E5.pdf: “poder identificar algunas que cosas que podríamos estar haciendo mal y que provocan sentimientos negativos en nosotros y aquellas que nos causan el no tener una motivación para estudiar; y como consecuencia de esto afectar nuestro rendimiento en la escuela. Un ejemplo podía ser el estresarse por un examen, o por sacar una mala nota en este mismo.”

6:1 p 3 en E6.pdf: “aprendimos a reconocer los factores emocionales que afectan de manera negativa nuestro rendimiento académico, y como estos pueden ser canalizados hacia una meta de aprendizaje mediante la motivación intrínseca y la inteligencia emocional.”

### **Factores Motivacionales**

Los segmentos de datos asociados a este código son:

2:4 p 4 en E2.pdf: “consideramos que la motivación es una parte importante para la vida universitaria ya que esta influye en nuestra salud mental. nos proporciona un punto de vista de cómo estas emociones afectan nuestros estudios.”

4:3 p 4 en E4.pdf: “nos ayudo saber e identificar las distintas maneras de motivación en las cuales podemos impulsarnos para alcanzar nuestros objetivos y/o metas a corto, mediano o largo plazo en cuanto nuestro desempeño en la carrera.”

4:4 p 4 en E4.pdf: “Como equipo consideramos la importancia de esta actividad ya que es indispensable que nosotros como estudiantes encontremos esos motivos por lo cual debemos esforzarnos para la consecución de los objetivos académicos y la mejora del rendimiento escolar.”

7:4 p 4 en E7.pdf: “nos dió un poco más de motivación, viendo lo que podemos hacer como licenciados en ciencias de la computación. Además de empezar a pensar en qué dirección queremos enfocar nuestra carrera.”

8:3 p 4 en E8.pdf: “supimos controlar y redireccionar esas malas motivaciones o sentimientos para un bien, ya no nos afectan tanto y nos permitió crecer como estudiantes.”

9:5 p 5 en E9.pdf: “los tipos de motivación que suelen tener los estudiantes en sus actividades, en teoría refiriéndose a la motivación extrínseca e intrínseca, con ellos se buscaba resaltar las principales diferencias y motivos por los que hacen presencia y de alguna manera mejorar dicha motivación o reforzarla.”

## **PI 2: ¿Qué tipos de conocimientos, habilidades o actitudes fueron favorecidas?**

Por cada una de las categorías de códigos se presentan fragmentos de textos que reflejan las percepciones de los estudiantes:

### **Autocontrol**

2:3 p 4 en E2.pdf “aprendimos a identificar los factores emocionales que nos afectan también aprendimos que la inteligencia emocional es más importante de lo que

creíamos ya que el aprender a controlar tus emociones y regularlas nos da un control de nosotros mismos en situaciones poco favorables para nosotros.”

4:6 p5 en E4.pdf “representa una forma de conocernos a nosotros mismos, así como el entender la forma en que nuestras emociones nos hacen propensos a no tomar buenas decisiones.”

7:1 p 2 en E7.pdf “Logramos un autocontrol de las distintas situaciones académicas y extraescolares que propician un bajo rendimiento académico por medio de la inteligencia emocional.”

8:3 p4 en E8.pdf “supimos controlar y redireccionar esas malas motivaciones o sentimientos para un bien, ya no nos afectan tanto y nos permitió crecer como estudiantes.”

### **Competencias digitales**

7:5 p2 en E7.pdf “Desarrollé más habilidades y conocimientos acerca de los distintos tipos de plataformas y entornos virtuales que me facilitan el aprendizaje y la comunicación en mi vida académica.”

8:2 p 3 en E8.pdf “logramos adquirir el conocimiento a cerca de los medios de comunicación y colaboración, como podemos aprovechar estas tecnologías a su máximo nivel para poder facilitarnos la realización de actividades a distancia.”

9:1 p 3 en E9.pdf “la necesidad de hablar acerca de las tecnologías en la actualidad, porque se han convertido en parte de nuestra vida cotidiana y a menudo van creciendo y evolucionando con nosotros.”

9:2 p 3 en E9.pdf “mejore mis habilidades para manejar grandes cantidades de información gracias al uso de las tecnologías como forma de organización y también gracias a recursos en línea para manejar situaciones complicadas como hoy en día con las clases en línea.”

## **Establecimiento de metas**

6:4 p 4 en E6.pdf “nos facilitó el establecimiento de metas y benefició la búsqueda de motivación para continuar con nuestros estudios de nivel superior.”

9:4 p 4 en E9.pdf “El conocer de que se trata la realidad de nuestro ámbito profesional para desarrollar el pensamiento crítico para el análisis de conceptos, teorías, explicaciones y puntos de vista para fomentar una idea de donde nos desenvolvemos mejor y crear una meta para prepararnos para ese camino.”

## **Estrategias de estudio**

2:7 p 9 en E2.pdf “Aprendimos estrategias de trabajo, individual y en equipo, para optimizar el tiempo y realizar el trabajo de manera más eficiente; por ejemplo, se pueden desmembrar los proyectos y tareas en secciones pequeñas y más manejables.”

4:9 p 11 en E4.pdf “Implementé estrategias de estudio que facilitan el aprendizaje del alumno y con las cuales logra una mejor comprensión y realización de la actividad.”

5:4 p 3 en E5.pdf “Lo que más aprendí fueron las diferentes técnicas que podemos aplicar para estudiar, me ayudó a poner en práctica nuevas formas de estudio.”

## **Hábitos de estudio**

3:2 p 4 en E3.pdf “Elegimos esta ADA por su contenido interesante de los hábitos que tienden a tener los estudiantes, ya que, entre ellos existen buenos y malos hábitos que tienen un efecto sobre nuestro desempeño académico en la universidad.”

3:3 p 4 en E3.pdf “nos ayudó a administrar mejor nuestro tiempo libre y de estudio, así como, descansar correctamente para poder ejercer un mejor desempeño. Pudimos controlar los malos hábitos que solemos siempre.”

6:7 p 8 en E6.pdf “aprendí a identificar como los buenos hábitos del estudio son beneficiosos y como en mi vida debo aplicarlos y afortunadamente lo estoy haciendo.”

### **Manejo de emociones**

5:1 p 3 en E5.pdf “poder identificar algunas que cosas que podríamos estar haciendo mal y que provocan sentimientos negativos en nosotros y aquellas que nos causan el no tener una motivación para estudiar; y como consecuencia de esto afectar nuestro rendimiento en la escuela. Un ejemplo podía ser el estresarse por un examen, o por sacar una mala nota en este mismo.”

6:1 p 3 en E6.pdf “aprendimos a reconocer los factores emocionales que afectan de manera negativa nuestro rendimiento académico, y como estos pueden ser canalizados hacia una meta de aprendizaje mediante la motivación intrínseca y la inteligencia emocional.”

2:10 p 8 en E2.pdf “Me conocí a mí misma y también a como usar las emociones a mi favor y no en contra, para motivarme a seguir mis objetivos evitando caer en lo que llamaría un agujero de depresión.”

### **Metacognición**

4:7 p 7 en E4.pdf “Conforme fue transcurriendo la unidad me volví mas consciente de lo que se encontraba a mi alrededor, todo lo que de alguna me influenciaba o me afectaba conforme a mi educación.”

6:5 p 6 en E6.pdf “A lo largo de esta unidad aprendí cuales son los factores que influyen en mi rendimiento académico.”

4:2 p 3 en E4.pdf “La competencia que desarrollamos al realizar esta actividad fue la de ejercer nuestro pensamiento crítico para analizar la realidad de nuestra carrera.”

2:8 p 10 en E2.pdf “Reflexioné mucho acerca de los valores y aptitudes que debo tener como egresado y estudiante de LCC, como serviré a la sociedad, los retos que debo afrontar y cómo debe ser mi perfil.”

### **Motivación**

3:6 p6 E3.pdf” se vio la importancia de la motivación en el aprendizaje y como puede ser afectada por las emociones y el entorno, logrando así que pueda ser tanto positiva como negativa.”

6:8 p 6 en E6.pdf “aprendí acerca de la motivación y el impacto e importancia que tiene en la vida universitaria de una persona.”

7:4 p 4 en E7.pdf “nos dió un poco más de motivación, viendo lo que podemos hacer como licenciados en ciencias de la computación. Además de empezar a pensar en qué dirección queremos enfocar nuestra carrera.”

### **Planificación**

7:2 p 3 en E7.pdf “Pensamos que es un punto de partida importante que nos ayudaría a definir y a organizar nuestras actividades escolares de manera responsable y así obtener un aprendizaje significativo.”

8:5 p 5 en E8.pdf “Aprendimos la importancia que tienen dichas competencias para nuestra formación y supimos como desarrollarlas en cada materia.”

2:9 p 10 en E2.pdf “me ha dado la introspección necesaria para proyectarse en el futuro, y ver los caminos que esta carrera pone ante mi, señalando puntos importantes de cómo debemos aspirar a ser; en cuanto a personal, ha sido una de las unidades que más dirección me ha dado durante la materia.”

## **Tipos de aprendizaje**

1:2 p 3 en E1.pdf “Esta actividad nos hizo ver que existen tipos de aprendizajes que no conocíamos sin embargo los implementamos, ahora ya sabemos que tipo de aprendizaje usar dependiendo la situación.”

2:6 p 5 en E2.pdf “sabíamos que existían diversas formas de estudio, pero en esta actividad pudimos encontrar la creemos que se adapta a nuestras capacidades y carencias, también en dicha actividad aprendimos a reconocer las diferentes estrategias del estudio.”

4:8 p 7 en E4.pdf “he logrado aprender más acerca de las diversas modalidades de aprendizaje y como estas afectan positiva o negativamente a cada estudiante.”

5:2 p 4 en E5.pdf “Aprendimos las características de los tipos de aprendizaje, cómo es que se pueden llevar a cabo en las escuelas, con ello poder identificar cual es el que nosotros utilizamos y a como sacar un mejor provecho de este mismo mediante las estrategias de estudio que se acoplan mejor a cada una.”

6:6 p 7 en E6.pdf “conocí los principales estilos de aprendizaje a través de los cuales el ser humano adulto adquiere conocimientos, así como sus características, similitudes, diferencias y técnicas de estudio apropiadas.”