

UNIVERSIDAD MIGUEL HERNÁNDEZ DE ELCHE

ESCUELA POLITÉCNICA SUPERIOR DE ELCHE

GRADO EN INGENIERÍA INFORMÁTICA EN
TECNOLOGÍAS DE LA INFORMACIÓN



Desarrollo y Evaluación de Algoritmos de
Clasificación para la Predicción del
Rendimiento Estudiantil

TRABAJO FIN DE GRADO

Septiembre-2024

AUTOR: Matteo Fecci Espiñeira

DIRECTOR/ES: Francisco Federico García Crespi

AGRADECIMIENTOS

Me gustaría agradecer a mi familia, cuyo apoyo incondicional ha sido fundamental a lo largo de estos años. A mis amigos, por estar siempre a mi lado y apoyarme en mis momentos más difíciles. A mis profesores, en especial a Francisco Federico, y a mis compañeros de trabajo. Sin su ayuda y aliento, nada de esto habría sido posible.



ÍNDICE

1. Introducción	1
1.1. Planteamiento del problema	2
1.2. Justificación del TFG	3
1.3. Objetivos Generales	4
1.3.1. Objetivos Específicos	5
1.4. Estado del Arte	5
1.5. Estructura del tfg	7
2. Machine Learning	9
2.1. Introducción al Machine Learning	9
2.2. Tipos de Aprendizaje en Machine Learning	12
2.3. Técnicas y Algoritmos Comunes	14
2.4. Ingeniería de Características y Preprocesamiento de Datos	16
2.5. Evaluación y Validación de Modelos	19
2.6. Regularización y Prevención del Sobreajuste	21
2.7. Infraestructura y Herramientas para Machine Learning	24
2.8. Implicaciones Éticas del Uso de Algoritmos en el Ámbito Educativo .	27
3. Minería de datos	31
3.1. Introducción a la Minería de Datos	31

3.2. Técnicas de Minería de Datos	33
3.3. Herramientas y Plataformas para Minería de Datos	35
3.4. Aplicaciones de la Minería de Datos	36
3.5. Desafíos y Tendencias Futuras en Minería de Datos	38
4. Inteligencia Artificial	40
4.1. Introducción a la Inteligencia Artificial	40
4.2. Historia de la Inteligencia Artificial	42
4.3. Modelos de Inteligencia Artificial	45
4.4. Aplicaciones Actuales de la Inteligencia Artificial	47
4.5. Problemas Legales y Morales del Uso de la Inteligencia Artificial	49
5. Learning analytics	51
5.1. Introducción y Fundamentos Teóricos de Learning Analytics	51
5.2. Metodologías, Técnicas y Aplicaciones de Learning Analytics	53
5.3. Desafíos, Consideraciones Éticas y Tendencias Futuras en Learning Analytics	55
6. Metodología	57
6.1. Herramientas utilizadas	57
6.1.1. Python	57
6.1.2. Visual Studio Code (VS Code)	57
6.1.3. Kaggle	58

6.1.4. ChatGPT	58
6.2. Recopilación y preprocesamiento de datos	58
6.2.1. Dataset of Students' Performance Using Student Information System, Moodle and the Mobile Application 'eDify'	59
6.2.1.1. Ventajas del Dataset	62
6.2.1.2. Desventajas del Dataset	63
6.2.1.3. Mejoras Potenciales	63
6.2.2. Open University	64
6.2.2.1. Recolección de Datos	64
6.2.2.2. Descripción del Dataset	65
6.2.2.3. Distribución del Número de Clics por Estudiante	66
6.2.2.4. Distribución de Estudiantes por Género	67
6.2.2.5. Distribución de Estudiantes por Grupo de Edad	68
6.2.2.6. Número de Aprobados y Suspendidos	69
6.2.2.7. Preprocesamiento de Datos	69
6.2.2.8. Ventajas del Dataset	70
6.2.2.9. Desventajas del Dataset	70
6.2.2.10. Mejoras Potenciales	70
6.3. Implementación de Algoritmos	71
6.3.1. Entorno y Herramientas	71
6.3.2. Lectura de Datos	71

6.3.3.	RandomForest	72
6.3.4.	J48	72
6.3.5.	CART	72
6.3.6.	LogisticRegression	72
7.	Resultados	74
7.1.	RandomForest	75
7.1.1.	Resultados del entorno local	75
7.1.2.	Resultados del entorno apoyado por IA	77
7.1.3.	Conclusiones de la comparación	79
7.2.	J48	79
7.2.1.	Resultados del entorno local	79
7.2.2.	Resultados del entorno apoyado por IA	82
7.2.3.	Conclusiones de la comparación	84
7.3.	LogisticRegression	85
7.3.1.	Resultados del entorno local	85
7.3.2.	Resultados del entorno apoyado por IA	87
7.3.3.	Conclusiones de la comparación	89
7.4.	CART	90
7.4.1.	Resultados del entorno local	90
7.4.2.	Resultados del entorno apoyado por IA	93
7.4.3.	Conclusiones de la comparación	95

7.5. Comparación de Precisión	97
7.5.1. Resultados del entorno local	97
7.5.2. Resultados del entorno apoyado por IA	98
7.5.3. Comparación entre el entorno de IA y el entorno local	99
7.6. Conclusiones	100
7.7. Recomendaciones para futuras investigaciones	103
BIBLIOGRAFÍA	106



ÍNDICE DE FIGURAS

2.1. Histograma del Machine Learning (ML)[22]	10
2.2. Tipos de aprendizaje en ML	12
6.1. Proporción de Estudiantes Aprobados y Suspendidos	60
6.2. Distribución De Las Actividades En Línea Realizadas Dentro Del Campus	61
6.3. Distribución De Las Actividades En Línea Realizadas Fuera Del Cam- pus	61
6.4. Distribución del Número de Clics por Estudiante	66
6.5. Distribución de Estudiantes por Género	67
6.6. Distribución de Estudiantes por Grupo de Edad	68
6.7. Número de Aprobados y Suspendidos	69
7.1. Comparación de Matrices de Confusión para RandomForest.	75
7.2. Comparación de Curvas ROC para RandomForest.	76
7.3. Comparación de Matrices de Confusión en el entorno apoyado por IA.	77
7.4. Comparación de Curvas ROC en el entorno apoyado por IA.	78
7.5. Comparación de Matrices de Confusión para J48.	80
7.6. Comparación de Curvas ROC para J48.	81
7.7. Comparación de Matrices de Confusión en el entorno apoyado por IA.	82
7.8. Comparación de Curvas ROC en el entorno apoyado por IA.	83
7.9. Comparación de Matrices de Confusión para Regresión Logística.	85

7.10. Comparación de Curvas ROC para Regresión Logística.	86
7.11. Comparación de Matrices de Confusión en el entorno apoyado por IA.	87
7.12. Comparación de Curvas ROC en el entorno apoyado por IA.	88
7.13. Comparación de Matrices de Confusión para CART.	91
7.14. Comparación de Curvas ROC para CART.	92
7.15. Comparación de Matrices de Confusión en el entorno apoyado por IA.	93
7.16. Comparación de Curvas ROC en el entorno apoyado por IA.	94
7.17. Comparación de Precisión entre los Resultados A y B.	97
7.18. Comparación de Precisión entre los Modelos en el Entorno de IA.	98



ÍNDICE DE CÓDIGOS

ML Machine Learning

LA Learning Analytics

MD Minería de datos

IA Inteligencia artificial

MA Modelos de Aprendizaje

AP Análisis Predictivo

PLD Procesamiento de Lenguaje Natural (en inglés NLP)

GDPR Reglamento General de Protección de Datos

OECD Organización para la Cooperación y el Desarrollo Económico



1. Introducción

El sistema educativo ha evolucionado significativamente a lo largo de la historia, adaptándose a las transformaciones tecnológicas, sociales y económicas. En la era actual, la tecnología se ha convertido en una herramienta esencial para mejorar la calidad educativa y el proceso de enseñanza-aprendizaje. Entre las innovaciones tecnológicas más prometedoras se encuentra la aplicación del Machine Learning (ML) y la Minería de datos (MD) para la predicción del rendimiento estudiantil [1].

El ML se fundamenta en la capacidad de las máquinas para aprender a partir de datos y mejorar su rendimiento sin necesidad de programación explícita para cada tarea específica. En la educación, el ML se puede utilizar para analizar volúmenes significativos de datos generados por los estudiantes y extraer patrones que pueden ser empleados para optimizar la calidad de la enseñanza y personalizar el aprendizaje.

Una de las aplicaciones más prometedoras del ML en la educación es la predicción del rendimiento académico [1]. Esta técnica permite anticipar el desempeño de los estudiantes en función de una diversidad de factores, incluyendo su historial académico, participación en clase, hábitos de estudio y características personales, tales como la edad, el género o la ubicación. La identificación temprana de estudiantes en riesgo de bajo rendimiento permite a los educadores intervenir de manera oportuna y proporcionar el apoyo necesario para reforzar su aprendizaje.

A pesar de los numerosos beneficios que ofrece el ML en la educación, su implementación no está exenta de desafíos. Entre los principales retos se encuentran:

- **Privacidad y protección de datos:** La utilización de datos personales de los estudiantes para el entrenamiento de modelos de ML suscita importantes preocupaciones sobre la privacidad y la protección de datos. Es imperativo garantizar que el manejo de los datos se realice de manera ética y en cumplimiento con las normativas de protección de datos vigentes, como el Reglamento General de Protección de Datos (GDPR) en Europa [2].

- **Equidad y sesgo en los modelos:** Los modelos de ML pueden estar sujetos a sesgos inherentes en los datos utilizados para su entrenamiento. Si los datos de entrenamiento reflejan desigualdades o prejuicios existentes, los modelos pueden perpetuar o incluso exacerbar estas desigualdades. Es crucial desarrollar técnicas para la detección y mitigación del sesgo en los modelos predictivos [3].
- **Adopción y aceptación por parte de los educadores:** La integración de tecnologías avanzadas como el ML en el sistema educativo requiere la aceptación y el apoyo de los educadores. Es fundamental proporcionar formación y recursos adecuados para que los docentes comprendan cómo utilizar estas herramientas de manera efectiva y confíen en los beneficios que pueden aportar [4].
- **Infraestructura tecnológica:** La implementación de soluciones de ML en la educación requiere una infraestructura tecnológica adecuada, incluyendo acceso a equipos informáticos, software especializado y conectividad a internet. En muchos casos, las instituciones educativas pueden enfrentar limitaciones presupuestarias que dificultan la adopción de estas tecnologías [5].

1.1. Planteamiento del problema

La educación se encuentra en un proceso de cambio, impulsado por el desarrollo de las tecnologías digitales y el acceso cada vez mayor a los datos. Por eso, la capacidad de predecir los resultados académicos de los estudiantes mediante técnicas de ML se ha convertido en una herramienta valiosa para mejorar la calidad educativa [6], [7].

Para predecir los resultados académicos, hace falta analizar una variedad de factores que influyen en el rendimiento estudiantil, como características demográficas, hábitos de estudio, participación en actividades extracurriculares y antecedentes académicos. Este proceso tiene varias dificultades, como la recopilación y calidad de los datos, la identificación de variables relevantes y la complejidad de los modelos predictivos [8].

Uno de los problemas más significativos en Learning Analytics (LA) es la selección del algoritmo de ML adecuado. Con la aparición de técnicas nuevas y la mejora de las antiguas, elegir el algoritmo más efectivo se ha convertido en una tarea compleja [9]. El teorema de “No Free Lunch“ para la optimización establece que no existe un único algoritmo que sea superior en todos los escenarios, lo que implica que la efectividad de un algoritmo depende de las características específicas de los datos y del problema a resolver [9].

La evolución de los estos modelos predictivos ha permitido una mayor precisión en las predicciones, pero al mismo tiempo ha incrementado la complejidad en la selección del algoritmo más adecuado. La proliferación de algoritmos de ML ofrece una amplia gama de opciones, pero también exige una evaluación cuidadosa para determinar cuál de ellos se adapta mejor a las necesidades específicas de cada institución educativa[8].

1.2. Justificación del TFG

En el entorno educativo actual, la capacidad de predecir el rendimiento académico de los estudiantes se ha convertido en una herramienta valiosa para mejorar la calidad educativa. Las instituciones enfrentan el desafío de atender a una gran diversidad de estudiantes con recursos limitados, lo que exige la implementación de tecnologías que permitan tomar decisiones basadas en datos.

El uso de Machine Learning (ML) en el ámbito educativo responde a esta necesidad, proporcionando una oportunidad para anticipar el rendimiento académico a partir de una serie de factores. Esto incluye no solo el historial académico de los estudiantes, sino también su participación en clase, hábitos de estudio, características personales como edad y género, y otros datos demográficos. Este enfoque permite identificar tempranamente a aquellos estudiantes en riesgo de bajo rendimiento, ofreciendo a los educadores la posibilidad de intervenir de manera oportuna.

La precisión de las predicciones basadas en ML facilita la personalización del apoyo académico, optimizando la asignación de recursos. Por ejemplo, se puede gestionar mejor el tiempo de los docentes, asignando atención especial a los estudiantes con

más dificultades, y diseñar estrategias de intervención adecuadas para mejorar los resultados académicos. Además, esta tecnología tiene el potencial de reducir significativamente las tasas de abandono escolar, beneficiando tanto a los estudiantes como a las instituciones educativas.

1.3. Objetivos Generales

El objetivo principal de este trabajo es identificar y evaluar la eficacia de diversos algoritmos de ML en la predicción del rendimiento académico de los estudiantes, utilizando tanto un entorno apoyado en python generado por el alumno como un entorno local apoyado en Inteligencia artificial (IA) y supervisado por el alumno. Este estudio no solo busca determinar la precisión y fiabilidad de diferentes modelos, sino también comprender las condiciones bajo las cuales cada algoritmo opera de manera óptima. La comparación exhaustiva de estos algoritmos tiene como propósito proporcionar una base sólida de conocimientos que facilite el desarrollo futuro de herramientas predictivas avanzadas en el ámbito educativo.

Adicionalmente, este trabajo persigue un objetivo general más amplio: establecer las bases para el desarrollo de una herramienta de predicción del rendimiento académico utilizando técnicas de ML. Esta herramienta está concebida como una innovación significativa en el ámbito educativo, proporcionando a las instituciones educativas una capacidad avanzada para anticipar el desempeño de los estudiantes. La predicción futura del rendimiento académico permitirá a los educadores identificar tempranamente a los estudiantes que requieren apoyo adicional, diseñar intervenciones personalizadas y optimizar las estrategias de enseñanza. Este enfoque proactivo tiene el potencial de reducir las tasas de abandono escolar, aumentar las tasas de éxito académico y, en última instancia, contribuir a la creación de un entorno educativo más equitativo y eficiente.

El desarrollo de esta herramienta predictiva se fundamenta en una comprensión profunda de las variables y factores que influyen en el rendimiento académico. La identificación y análisis de estos factores no solo busca mejorar la precisión de las predicciones, sino también proporcionar insights valiosos que puedan orientar a los

educadores en la toma de decisiones informadas y basadas en datos. Este enfoque integral y detallado asegura que la herramienta de predicción no solo sea técnicamente robusta, sino también relevante y útil en el contexto educativo real.

1.3.1. Objetivos Específicos

Para alcanzar el objetivo general previamente establecido, se plantean los siguientes objetivos específicos:

- Implementar y evaluar una serie de algoritmos de clasificación ampliamente reconocidos en el campo del aprendizaje automático, incluyendo Random Forest, J48, Regresión Logística y CART[11].
- Realizar un análisis comparativo del rendimiento de los algoritmos implementados, utilizando métricas de evaluación estandarizadas tales como la precisión, las matrices de confusión y el área bajo la curva ROC (AUC). Este análisis tiene como objetivo determinar qué algoritmo proporciona las predicciones más precisas y fiables, facilitando así el desarrollo futuro de herramientas predictivas en el ámbito educativo.
- Desarrollar y contrastar dos enfoques de integración: una implementación local y una implementación utilizando herramientas de IA. Este objetivo busca evaluar la viabilidad y eficacia de diferentes entornos de desarrollo para la predicción del rendimiento académico.
- Implementar técnicas de validación cruzada, de las que hablaremos posteriormente, para garantizar la estabilidad y capacidad de generalización del modelo predictivo. Asimismo, se realizará un análisis detallado de los errores para mejorar iterativamente el modelo y reducir la tasa de errores de predicción.

1.4. Estado del Arte

El estado del arte en la predicción del rendimiento académico mediante ML y MD ha experimentado avances significativos en las últimas décadas. Este campo inter-

disciplinario combina técnicas avanzadas de análisis de datos, algoritmos de ML y conocimientos pedagógicos con el objetivo de optimizar los procesos educativos.

La predicción del rendimiento académico constituye una de las aplicaciones más estudiadas del ML en la educación. Diversos estudios han empleado modelos de ML para anticipar el éxito académico de los estudiantes basándose en una variedad de factores, incluyendo su historial académico, participación en clase, hábitos de estudio y características demográficas [12].

Entre los modelos más frecuentemente utilizados para la predicción del rendimiento académico se encuentran los árboles de decisión, los bosques aleatorios (random forests), las máquinas de soporte vectorial y las redes neuronales. Estos modelos han demostrado su eficacia en la identificación de patrones complejos en los datos educativos [13].

Un estudio destacado realizado por Yu y DiGangi [14] implementó un modelo de árboles de decisión para predecir el rendimiento académico de estudiantes universitarios. Los resultados evidenciaron una alta precisión en la predicción de las calificaciones finales, permitiendo a los educadores intervenir de manera temprana con los estudiantes en situación de riesgo académico.

La personalización del aprendizaje constituye otra área clave donde el ML ha tenido un impacto significativo. Mediante el análisis del rendimiento y las preferencias de los estudiantes, los algoritmos de ML pueden adaptar el contenido educativo y los métodos de enseñanza a las necesidades individuales de cada estudiante [15].

Las plataformas de aprendizaje adaptativo utilizan algoritmos de ML para proporcionar recomendaciones personalizadas y ajustar el nivel de dificultad de las tareas educativas en tiempo real. Un ejemplo destacado es la plataforma Knewton, que emplea aprendizaje profundo para personalizar la experiencia de aprendizaje de millones de estudiantes [16].

El análisis de sentimientos y la monitorización de la participación de los estudiantes son aplicaciones emergentes del ML en la educación. Mediante el análisis de los datos de participación en clase, como las interacciones en plataformas de aprendizaje

en línea, los algoritmos de ML pueden evaluar el nivel de compromiso y detectar problemas potenciales [17].

En un estudio realizado por Wen et al. [18], se utilizó el análisis de sentimientos para evaluar las respuestas de los estudiantes en foros de discusión en línea. Los resultados permitieron identificar a los estudiantes que estaban experimentando dificultades y proporcionar apoyo adicional.

1.5. Estructura del tfg

El marco teórico de este trabajo se desarrolla a lo largo de los capítulos 2, 3, 4 y 5. En estos capítulos se realiza una revisión de los conceptos fundamentales del ML, MD, IA, y se presenta un estado del arte actualizado del LA.

El capítulo 6 consiste en la descripción de los entornos que se han usado para el trabajo, facilitando así la reproducción y mejora del mismo.

En el capítulo 7 se expone la metodología seguida en la investigación. Esta sección detalla los procedimientos, técnicas y enfoques adoptados para abordar los objetivos planteados en el estudio.

Los capítulos subsiguientes corresponden al desarrollo del trabajo en sí, presentando los resultados obtenidos, su análisis e interpretación. El documento culmina con un capítulo dedicado a las conclusiones, donde se sintetizan los hallazgos principales y se discuten sus implicaciones para el campo de la educación asistida por ML.

Finalmente, se incluye una sección bibliográfica, que recoge todas las fuentes consultadas y citadas a lo largo del trabajo.

2. Machine Learning

2.1. Introducción al Machine Learning

El ML, una subdisciplina de la IA, se define como un campo que utiliza algoritmos estadísticos y computacionales para mejorar su desempeño en una tarea específica mediante la experiencia. Este proceso de mejora continua se denomina "aprendizaje". Los principales conceptos fundamentales en ML incluyen el algoritmo, que es un conjunto de reglas y procedimientos que un modelo de ML sigue para aprender de los datos; el modelo, que es la representación matemática del sistema que se entrena con datos; los datos de entrenamiento, que son el dataset que se utiliza para enseñar al modelo; la función de costo, que es una función que el algoritmo de aprendizaje trata de minimizar durante el entrenamiento; y la generalización, que es la capacidad del modelo para aplicar lo aprendido a nuevos datos no vistos anteriormente.[19]

Historia del Machine Learning: El campo del ML ha evolucionado significativamente desde sus inicios en la década de 1950. En sus primeros días, los investigadores se centraban en el desarrollo de algoritmos básicos y en la comprensión de los principios fundamentales del aprendizaje automático. Con el tiempo, el aumento de la capacidad de procesamiento de las computadoras y la disponibilidad de grandes volúmenes de datos impulsaron avances rápidos en el campo. En las décadas de 1980 y 1990, se desarrollaron técnicas clave como las redes neuronales y las máquinas de vectores de soporte (SVM), que se convirtieron en herramientas fundamentales para la clasificación y regresión [20], [21].

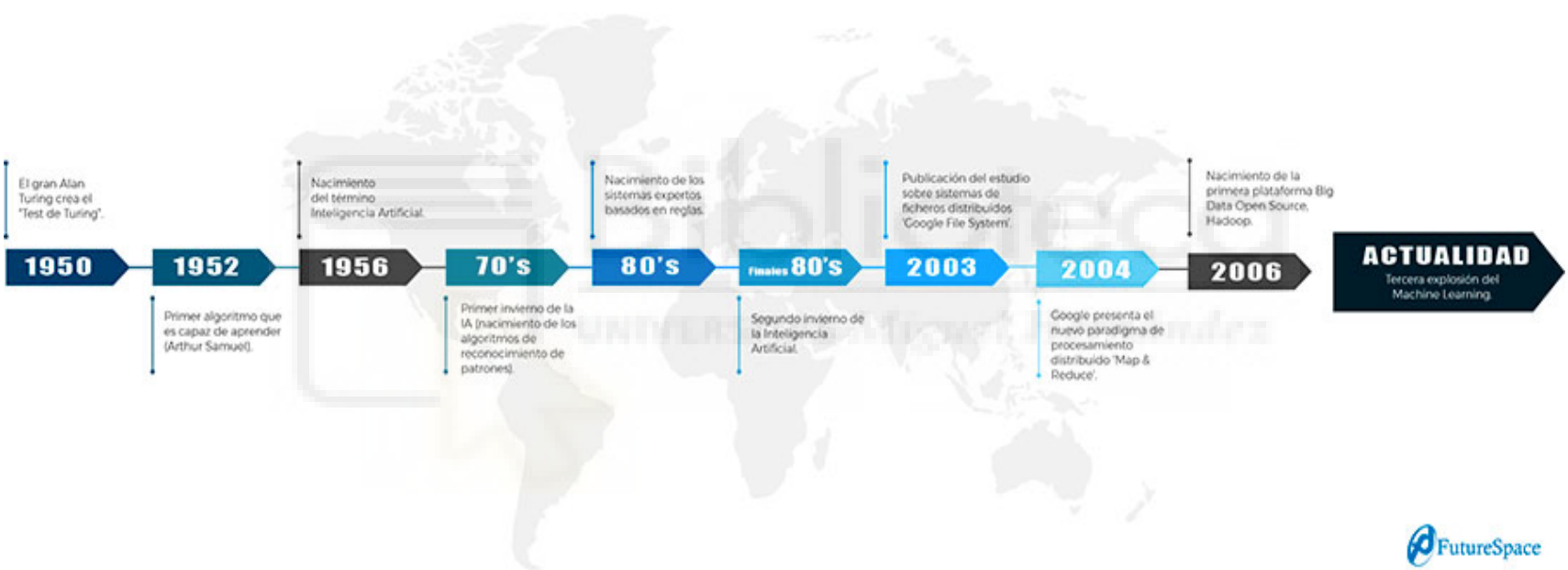


Figura 2.1: Histograma del ML[22]

En la figura 2.1 podemos ver, de forma muy esquemática, la evolución del ML a lo largo de la historia, desde sus inicios hasta la actualidad, encontrándonos actualmente en la tercera explosión del ML. En esta explosión, los avances se empiezan a usar, de manera significativas, en el mundo empresarial, llegando hasta el punto de crear completos mercados y produciendo cambios significativos en las estrategias de grandes empresas.

El ML moderno se caracteriza por el uso de grandes conjuntos de datos y potentes algoritmos de aprendizaje para abordar problemas complejos. Los frameworks y bibliotecas de ML, como TensorFlow, PyTorch y Scikit-Learn, han facilitado enormemente el desarrollo y la implementación de modelos sofisticados, acelerando la adopción de ML en la industria y la academia [23], [24].

Aplicaciones del Machine Learning: El ML se aplica en una amplia gama de campos y ha transformado numerosas industrias. Algunos de los principales ámbitos de aplicación incluyen la visión por computadora, utilizada en la identificación y clasificación de objetos en imágenes y videos; el Procesamiento de Lenguaje Natural (en inglés NLP) (PLD), aplicado en tareas como la traducción automática, la generación de texto y el análisis de sentimientos; los sistemas de recomendación, implementados por servicios como Netflix y Amazon para sugerir productos o contenido a los usuarios basándose en sus preferencias y comportamientos pasados; el diagnóstico médico, que ayuda en la detección y diagnóstico de enfermedades a partir de datos médicos y de imágenes, mejorando la precisión y rapidez de los diagnósticos y facilitando la medicina personalizada; y las finanzas, donde se usa en la predicción de precios de acciones, detección de fraudes y gestión de riesgos [25].

A pesar de sus numerosos beneficios, el ML enfrenta varios desafíos y consideraciones importantes. La calidad y cantidad de los datos utilizados para el entrenamiento de los modelos son cruciales para su desempeño, por lo que datos incompletos o sesgados pueden llevar a modelos ineficaces o incluso perjudiciales, haciendo de la recopilación y preprocesamiento de datos etapas críticas en cualquier proyecto de ML. La interpretabilidad de los modelos es un aspecto clave, especialmente en áreas donde la transparencia es crucial, como en la medicina y las finanzas. Métodos como SHAP (SHapley Additive exPlanations) y LIME (Local Interpretable Model-

agnostic Explanations) se utilizan para explicar las predicciones de los modelos, proporcionando una mejor comprensión de los factores que influyen en las decisiones del modelo [26], [27].

Las consideraciones éticas y la mitigación de sesgos son fundamentales en el desarrollo de sistemas de ML. Los algoritmos de ML pueden perpetuar o amplificar sesgos existentes en los datos de entrenamiento, lo que puede llevar a decisiones injustas o discriminatorias. Es esencial que los desarrolladores adopten un enfoque consciente y crítico al diseñar y evaluar modelos, asegurando que se identifiquen y mitiguen los posibles sesgos [28].

2.2. Tipos de Aprendizaje en Machine Learning

El ML se puede clasificar principalmente en tres tipos de aprendizaje: supervisado, no supervisado y por refuerzo. Cada uno de estos tipos tiene características y aplicaciones únicas que los hacen adecuados para diferentes tipos de problemas y datos [29].

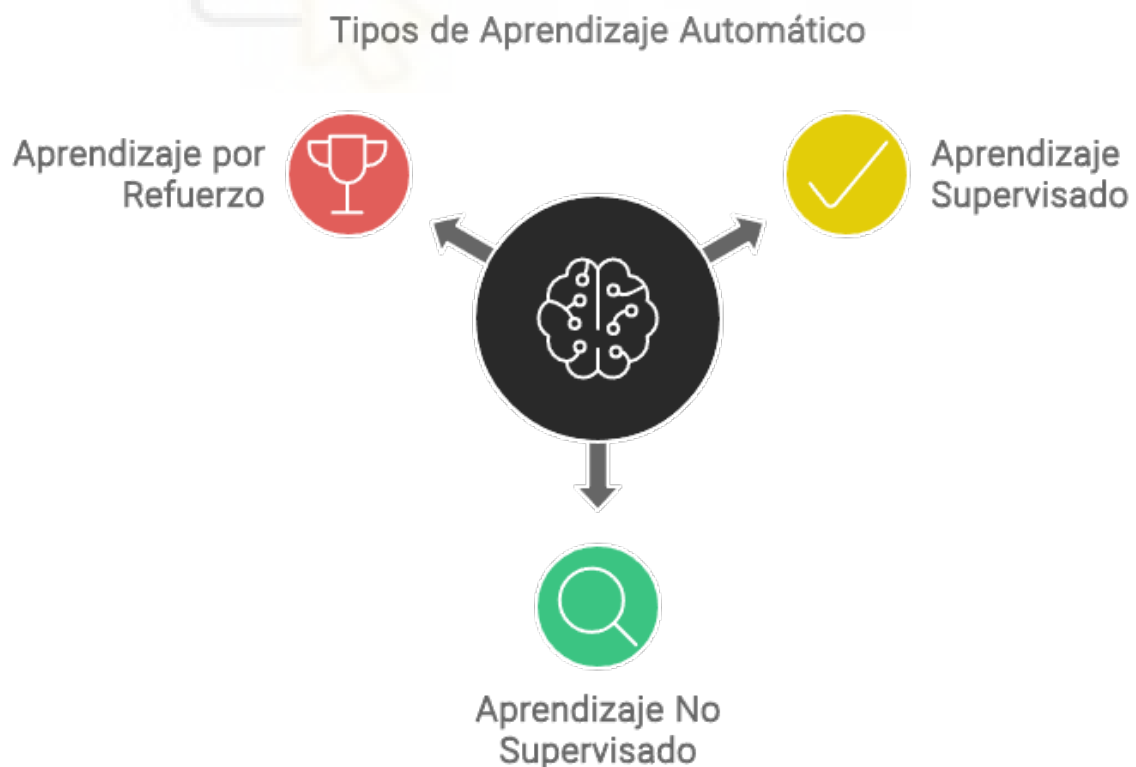


Figura 2.2: Tipos de aprendizaje en ML

El **aprendizaje supervisado** es, quizás, el tipo más común de ML. En este enfoque, el modelo se entrena con un dataset etiquetado, lo que significa que cada entrada de datos está asociada con una salida correcta. El objetivo es que el modelo aprenda a mapear entradas a salidas correctas para predecir resultados de nuevos datos. Los algoritmos de aprendizaje supervisado pueden ser divididos en dos categorías principales: clasificación y regresión [30].

En los problemas de clasificación, el objetivo es asignar una etiqueta a cada instancia de datos. Ejemplos de algoritmos de clasificación incluyen la regresión logística, las máquinas de vectores de soporte (SVM), y los árboles de decisión [31]. Por ejemplo, en la detección de spam, el objetivo es clasificar los correos electrónicos como "spam" o "correo legítimo". En los problemas de regresión, el objetivo es predecir un valor continuo. Ejemplos de algoritmos de regresión incluyen la regresión lineal y las redes neuronales. Un ejemplo típico sería predecir los precios de la vivienda basándose en características como el tamaño, la ubicación y el número de habitaciones [32].

El **aprendizaje no supervisado** se utiliza cuando los datos no están etiquetados y el objetivo es identificar patrones ocultos o estructuras en los datos. Este tipo de aprendizaje es útil para tareas como la agrupación (clustering) y la reducción de dimensionalidad. En la agrupación, el objetivo es dividir los datos en grupos (clusters) de manera que las instancias dentro de cada grupo sean más similares entre sí que a las de otros grupos. Algoritmos como K-means y DBSCAN son ejemplos de técnicas de agrupación (clustering)[29]. En la reducción de dimensionalidad, el objetivo es reducir el número de variables en los datos mientras se preserva la mayor cantidad de información posible. El Análisis de Componentes Principales (PCA) y el t-SNE son técnicas comunes utilizadas para este propósito [32]. La reducción de dimensionalidad es particularmente útil en la visualización de datos y en la preprocesamiento de datos para otros algoritmos de ML.

El **aprendizaje por refuerzo** es un tipo de aprendizaje que se basa en la idea de que un agente aprende a comportarse en un entorno realizando acciones y recibiendo recompensas o castigos. El objetivo es maximizar la recompensa acumulada a lo largo del tiempo. Este tipo de aprendizaje es particularmente relevante en áreas como la robótica, los juegos y la conducción autónoma [33]. En el aprendizaje por refuerzo, el

agente toma decisiones basadas en un estado del entorno y recibe una recompensa basada en la acción realizada. Algoritmos como Q-Learning, SARSA y las Redes Neuronales Profundas de Aprendizaje por Refuerzo (Deep Q-Networks, DQN) son populares en este campo [34].

Cada uno de estos tipos de aprendizaje tiene sus propias ventajas y desafíos. El aprendizaje supervisado suele ser más preciso ya que se entrena con ejemplos etiquetados, pero puede requerir grandes cantidades de datos etiquetados, lo cual puede ser costoso y laborioso de obtener [29]. El aprendizaje no supervisado puede descubrir patrones ocultos en los datos, pero los resultados pueden ser menos interpretables y más difíciles de evaluar [30]. El aprendizaje por refuerzo puede adaptarse dinámicamente a entornos cambiantes y aprender estrategias complejas, pero requiere un diseño cuidadoso de la función de recompensa y puede tener una alta carga computacional [33].

La elección del tipo de aprendizaje adecuado depende del problema específico que se esté abordando, la disponibilidad de datos etiquetados y los recursos computacionales disponibles. En muchos casos, una combinación de diferentes tipos de aprendizaje puede ser la mejor estrategia para obtener resultados óptimos. Por ejemplo, en un sistema de recomendación, se pueden utilizar técnicas de aprendizaje supervisado para predecir la calificación de un usuario para un producto específico, técnicas de aprendizaje no supervisado para agrupar a usuarios con preferencias similares, y técnicas de aprendizaje por refuerzo para optimizar la estrategia de recomendación en tiempo real [35].

2.3. Técnicas y Algoritmos Comunes

En ML se usan una variedad de técnicas y algoritmos, cada uno adecuado para diferentes tipos de problemas y datos. A continuación, se describen algunos de los algoritmos más comunes utilizados en la práctica del ML.

Regresión Lineal: La regresión lineal es una técnica simple para modelar la relación entre una variable dependiente y una o más variables independientes. Se utiliza para

predecir un valor numérico continuo basado en la ecuación de una línea recta. Es particularmente útil en problemas de predicción y análisis de tendencias. La ecuación general de una regresión lineal simple es $y = \beta_0 + \beta_1x$, donde y es la variable dependiente, x es la variable independiente, y β_0 y β_1 son los coeficientes que se ajustan durante el entrenamiento [36].

Regresión Logística: Utilizada para modelos de clasificación binaria, la regresión logística predice la probabilidad de que una instancia pertenezca a una clase determinada. A diferencia de la regresión lineal, su salida es una probabilidad que se convierte en una clasificación mediante un umbral. La ecuación de la regresión logística se basa en la función logística o sigmoide, que transforma la salida de una combinación lineal de las características en un valor entre 0 y 1 [37].

Máquinas de Vectores de Soporte (SVM): Las SVM son algoritmos de clasificación que buscan un hiperplano que mejor separe las clases en los datos. Este hiperplano es aquel que maximiza la distancia (margen) entre los puntos de datos de las dos clases más cercanos al hiperplano, conocidos como vectores de soporte. Las SVM son particularmente efectivas en problemas de alta dimensionalidad y se pueden extender a problemas de regresión [21].

Redes Neuronales: Inspirados en la estructura del cerebro humano, las redes neuronales consisten en capas de neuronas artificiales conectadas entre sí. Estas capas incluyen una capa de entrada, una o más capas ocultas y una capa de salida. Las redes neuronales son capaces de aprender representaciones complejas y no lineales de los datos. Con la llegada del aprendizaje profundo (Deep Learning), las redes neuronales profundas (DNN) han demostrado ser extremadamente efectivas en tareas como el reconocimiento de imágenes, el PLD y la conducción autónoma [19].

Árboles de Decisión: Los árboles de decisión son algoritmos que dividen los datos en ramas basadas en características hasta alcanzar una predicción final. Cada nodo interno representa una prueba en una característica, cada rama representa el resultado de la prueba, y cada nodo hoja representa una clase o un valor de predicción. Los árboles de decisión son fáciles de interpretar y visualizar, pero pueden ser propensos al sobreajuste [38].

Bosques Aleatorios: Un bosque aleatorio es un conjunto de árboles de decisión entrenados con diferentes subconjuntos de datos y características. La predicción final se obtiene mediante la agregación (por ejemplo, promedio o votación mayoritaria) de las predicciones de los árboles individuales. Este enfoque reduce el sobreajuste y mejora la precisión [39].

K-Nearest Neighbors (KNN): El algoritmo KNN es un método de clasificación basado en la proximidad de una instancia a sus vecinos más cercanos en el espacio de características. Para predecir la clase de una instancia nueva, el algoritmo considera las K instancias más cercanas en los datos de entrenamiento y asigna la clase más común entre ellas. KNN es simple y efectivo, pero puede ser computacionalmente costoso con grandes conjuntos de datos [40].

Cada uno de estos algoritmos tiene sus propias ventajas y limitaciones, y la elección del algoritmo adecuado depende del tipo de problema, la naturaleza de los datos y los requisitos específicos del proyecto. Además, a menudo se utilizan en combinación con técnicas de preprocesamiento de datos, selección de características y técnicas de validación cruzada para optimizar el rendimiento del modelo [31].

2.4. Ingeniería de Características y Preprocesamiento de Datos

La ingeniería de características y el preprocesamiento de datos son etapas críticas en cualquier proyecto de ML. Estas etapas implican la preparación de los datos en un formato adecuado para el análisis, mejorando la calidad de los datos y, en última instancia, el rendimiento del modelo. A continuación, se describen algunas de las técnicas más comunes en esta área

Selección y Extracción de Características: La selección de características implica elegir el subconjunto más relevante de las características originales para usar en el modelo. Esto puede ayudar a reducir la complejidad del modelo y mejorar su rendimiento. Las técnicas de selección de características incluyen métodos de filtrado, envoltura y basados en árboles. Por ejemplo, los métodos de filtrado utilizan prue-

bas estadísticas para evaluar la importancia de cada característica, mientras que los métodos de envoltura evalúan diferentes subconjuntos de características utilizando un modelo predictivo. Los métodos basados en árboles, como los bosques aleatorios, pueden medir la importancia de las características basándose en su contribución a la reducción de la impureza en los nodos [41].

La extracción de características crea nuevas características a partir de las originales, capturando la información más relevante. El Análisis de Componentes Principales (PCA) es un método popular para la extracción de características, que reduce la dimensionalidad del dataset mientras retiene la mayor variabilidad posible. PCA transforma las características originales en un nuevo conjunto de características no correlacionadas, conocidas como componentes principales, ordenadas por la cantidad de variabilidad que explican en los datos [42].

Normalización y Estandarización de Datos: La normalización y la estandarización son técnicas utilizadas para escalar las características a un rango similar, lo que puede mejorar la convergencia de los algoritmos de ML y evitar que las características con valores más grandes dominen el proceso de aprendizaje. La normalización escala los valores de las características a un rango de 0 a 1, lo cual es útil para algoritmos que dependen de la distancia, como K-Nearest Neighbors (KNN) y los algoritmos de clustering. La estandarización transforma los datos para que tengan una media de 0 y una desviación estándar de 1, lo que beneficia a algoritmos que asumen una distribución normal de los datos, como la regresión logística y las redes neuronales. [32].

Tratamiento de Datos Faltantes: Los datos faltantes son un problema común en los conjuntos de datos del mundo real. Existen varias técnicas para manejar los datos faltantes, incluyendo la eliminación de filas o columnas con valores faltantes, la imputación de valores faltantes utilizando la media, la mediana o la moda, y el uso de algoritmos más avanzados como la imputación mediante modelos predictivos. La elección de la técnica adecuada depende de la cantidad de datos faltantes y de la importancia de las características afectadas [43].

Codificación de Variables Categóricas: Las características categóricas deben

ser transformadas en un formato numérico antes de ser utilizadas en los modelos de ML. Las técnicas comunes para la codificación de variables categóricas incluyen la codificación one-hot, donde cada categoría se representa como una columna binaria separada, y la codificación ordinal, donde las categorías se asignan a valores enteros. La codificación one-hot es útil cuando no hay un orden inherente entre las categorías, mientras que la codificación ordinal es apropiada cuando las categorías tienen un orden natural [44].

Detección y Tratamiento de Valores Atípicos: Los valores atípicos pueden afectar negativamente el rendimiento del modelo. La detección de valores atípicos se puede realizar utilizando métodos estadísticos, gráficos de caja y bigotes, o técnicas más avanzadas como el análisis de componentes principales y los algoritmos de clustering. Una vez detectados, los valores atípicos pueden ser eliminados, transformados o imputados. Por ejemplo, los métodos basados en el rango intercuartílico identifican valores atípicos como aquellos que caen fuera de 1.5 veces el rango intercuartílico por encima del tercer cuartil o por debajo del primer cuartil [45].

Transformación de Características: En algunos casos, las transformaciones de características pueden ayudar a mejorar la linealidad de las relaciones entre las características y las etiquetas, o a manejar mejor la distribución de los datos. Las transformaciones comunes incluyen la transformación logarítmica, la transformación de raíz cuadrada y la normalización z-score. Por ejemplo, la transformación logarítmica puede ser útil para características con una distribución sesgada, convirtiendo una relación exponencial en una relación lineal [46].

La selección y transformación de características también pueden involucrar técnicas específicas del dominio. Por ejemplo, en problemas de visión por computadora, la extracción de características puede incluir la identificación de bordes, texturas y formas en las imágenes. En el PLD, las técnicas de ingeniería de características pueden incluir la tokenización, la eliminación de stopwords y la lematización. La adaptación de estas técnicas a las características específicas del problema puede mejorar significativamente el rendimiento del modelo.

El preprocesamiento de datos es un proceso iterativo que puede requerir múltiples

rondas de ajuste y validación. A medida que se desarrollan y prueban modelos, puede ser necesario volver a los datos para realizar ajustes adicionales en las características y el preprocesamiento. Este enfoque iterativo garantiza que el modelo final esté bien ajustado a los datos y sea capaz de generalizar a nuevos casos de manera efectiva.

2.5. Evaluación y Validación de Modelos

La evaluación y validación de modelos son etapas esenciales en el proceso de desarrollo de ML, ya que garantizan que el modelo funcione bien no solo en los datos de entrenamiento sino también en datos no vistos previamente. Estas etapas ayudan a identificar problemas como el sobreajuste y a seleccionar el modelo más adecuado para el problema en cuestión.

Métricas de Desempeño: La selección de métricas de desempeño adecuadas depende del tipo de problema que se esté abordando. Para problemas de clasificación, las métricas comunes incluyen la precisión, la sensibilidad (recall), la especificidad, el valor F1 y el área bajo la curva ROC (AUC-ROC). La precisión mide la proporción de predicciones correctas, mientras que la sensibilidad (recall) evalúa la capacidad del modelo para identificar las instancias positivas. La especificidad mide la capacidad del modelo para identificar correctamente las instancias negativas, y el valor F1 es la media armónica de la precisión y la sensibilidad, proporcionando un equilibrio entre ambas. La AUC-ROC es una medida del rendimiento del modelo en diferentes umbrales de clasificación [47].

Para problemas de regresión, las métricas típicas incluyen el error cuadrático medio (MSE), el error absoluto medio (MAE) y el coeficiente de determinación (R^2). El MSE mide la media de los cuadrados de los errores, penalizando más los errores grandes. El MAE mide la media de los valores absolutos de los errores, proporcionando una medida más directa de la magnitud del error. El R^2 indica la proporción de la variabilidad en la variable dependiente que es explicada por las variables independientes en el modelo [36].

Validación Cruzada: La validación cruzada es una técnica utilizada para evaluar

el desempeño de un modelo de manera robusta. Consiste en dividir los datos en múltiples subconjuntos y entrenar el modelo varias veces, cada vez utilizando un subconjunto diferente como conjunto de validación y los restantes como conjunto de entrenamiento. La validación cruzada más común es la validación cruzada k-fold, donde los datos se dividen en k partes iguales. El modelo se entrena k veces, cada vez utilizando k-1 partes para el entrenamiento y una parte diferente para la validación. Esto ayuda a garantizar que el modelo generalice bien a datos no vistos y no se ajuste en exceso a los datos de entrenamiento [48].

Conjunto de Validación y Conjunto de Prueba: En el desarrollo de modelos de ML, los datos generalmente se dividen en tres conjuntos: conjunto de entrenamiento, conjunto de validación y conjunto de prueba. El conjunto de entrenamiento se utiliza para ajustar los parámetros del modelo, el conjunto de validación se utiliza para ajustar los hiperparámetros y evaluar diferentes modelos, y el conjunto de prueba se utiliza para evaluar el desempeño final del modelo. Esta división ayuda a garantizar que el modelo se evalúe de manera justa y que su rendimiento refleje su capacidad para generalizar a nuevos datos [25].

Regularización y Prevención del Sobreajuste: El sobreajuste ocurre cuando un modelo se ajusta demasiado a los datos de entrenamiento, capturando ruido en lugar de patrones subyacentes. Esto resulta en un mal desempeño en datos no vistos. La regularización es una técnica utilizada para prevenir el sobreajuste añadiendo una penalización a la función de costo del modelo para evitar que aprenda patrones demasiado específicos de los datos de entrenamiento. Las formas comunes de regularización incluyen L1 (Lasso) y L2 (Ridge), que penalizan la magnitud de los coeficientes del modelo. La regularización L1 tiende a producir modelos más esparsos con menos características, mientras que la regularización L2 distribuye la penalización entre todas las características [49].

Técnicas de Evaluación Adicionales: Además de las métricas y técnicas mencionadas, existen otros métodos para evaluar el desempeño de los modelos. Las curvas de aprendizaje, por ejemplo, muestran el desempeño del modelo en función del tamaño del dataset de entrenamiento, ayudando a identificar si el modelo se beneficiaría de más datos. Los gráficos de error y las matrices de confusión son he-

herramientas visuales que permiten una comprensión más detallada de los errores del modelo y de cómo se distribuyen las predicciones incorrectas entre las diferentes clases [50].

La evaluación y validación de modelos son procesos iterativos que pueden requerir múltiples rondas de ajuste y pruebas. A medida que se desarrollan y prueban modelos, puede ser necesario volver a los datos para realizar ajustes adicionales en las características y el preprocesamiento, así como en los hiperparámetros del modelo. Este enfoque iterativo garantiza que el modelo final esté bien ajustado a los datos y sea capaz de generalizar a nuevos casos de manera efectiva .

Además, la implementación de modelos de ML en producción requiere un monitoreo continuo para garantizar que sigan funcionando correctamente y proporcionando valor en el entorno operativo. Esto puede incluir la implementación de sistemas de monitoreo para detectar y responder a cambios en el desempeño del modelo, así como la actualización periódica del modelo con nuevos datos para mantener su relevancia y precisión.

2.6. Regularización y Prevención del Sobreajuste

El sobreajuste es un problema común en el desarrollo de modelos de ML, donde un modelo se ajusta demasiado a los datos de entrenamiento y pierde su capacidad para generalizar a datos nuevos. La regularización es una técnica clave para prevenir el sobreajuste, añadiendo una penalización a la función de costo del modelo para evitar que aprenda patrones demasiado específicos de los datos de entrenamiento.

Regularización L1 (Lasso): La regularización L1 añade una penalización proporcional a la suma de los valores absolutos de los coeficientes del modelo. Esta penalización reduce algunos coeficientes a cero, lo que genera modelos más esparsos, que suelen ser más interpretables y menos propensos al sobreajuste. La función de costo para la regresión lineal con regularización L1 se define como:

$$\text{Costo} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^p |\beta_j|$$

donde n es el número de observaciones, y_i es el valor real, \hat{y}_i es el valor predicho, λ es el parámetro de regularización que controla la fuerza de la penalización, y β_j son los coeficientes del modelo [49].

Regularización L2 (Ridge): La regularización L2 añade una penalización proporcional a la suma de los cuadrados de los coeficientes del modelo. Esta penalización tiende a distribuir la penalización entre todas las características, reduciendo la magnitud de los coeficientes pero no llevándolos a cero. La función de costo para la regresión lineal con regularización L2 se define como:

$$\text{Costo} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^p \beta_j^2$$

Esta técnica ayuda a evitar que cualquier característica individual domine el modelo, mejorando su capacidad de generalización [51].

Regularización Elastic Net: Elastic Net es una técnica que combina las penalizaciones L1 y L2. Esto permite aprovechar las ventajas de ambas regularizaciones, promoviendo la esparsidad del modelo y controlando la magnitud de los coeficientes. La función de costo para Elastic Net se define como:

$$\text{Costo} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda_1 \sum_{j=1}^p |\beta_j| + \lambda_2 \sum_{j=1}^p \beta_j^2$$

donde λ_1 y λ_2 son los parámetros de regularización que controlan la fuerza de las penalizaciones L1 y L2, respectivamente [52].

Técnicas de Regularización para Redes Neuronales: En el contexto de las redes neuronales, además de las regularizaciones L1 y L2, se utilizan técnicas adicionales como el *dropout*. El *dropout* es una técnica en la que, durante el entrenamiento, se omiten aleatoriamente algunas neuronas de la red en cada iteración. Esto impide

que las neuronas desarrollen co-dependencias excesivas entre sí, mejorando la capacidad de generalización del modelo. En cada iteración, las neuronas que no se omiten se escalan para mantener la suma de activaciones constante [53].

Validación Cruzada: La validación cruzada es una técnica crucial no solo para evaluar el rendimiento del modelo, sino también para prevenir el sobreajuste. Al dividir los datos en múltiples subconjuntos y entrenar el modelo varias veces, cada vez utilizando un subconjunto diferente como conjunto de validación, se garantiza que el modelo generalice bien a datos no vistos. La validación cruzada k-fold es particularmente útil, ya que permite utilizar todos los datos para entrenamiento y validación en diferentes momentos, proporcionando una evaluación robusta del desempeño del modelo [48].

Early Stopping: Otra técnica utilizada para prevenir el sobreajuste es el *early stopping*. Durante el entrenamiento de modelos complejos como las redes neuronales, el rendimiento en el conjunto de validación puede comenzar a deteriorarse después de un cierto número de iteraciones, aunque el rendimiento en el conjunto de entrenamiento continúe mejorando. El *early stopping* detiene el entrenamiento cuando el rendimiento en el conjunto de validación deja de mejorar, evitando que el modelo se ajuste demasiado a los datos de entrenamiento [54].

Selección de Modelo: La selección del modelo adecuado también es crucial para prevenir el sobreajuste. Modelos más complejos tienen una mayor capacidad para aprender patrones detallados en los datos, pero también corren un mayor riesgo de sobreajuste. Modelos más simples pueden no capturar toda la complejidad de los datos, pero tienden a generalizar mejor. Herramientas como la validación cruzada y las curvas de aprendizaje pueden ayudar a identificar el modelo que ofrece el mejor equilibrio entre sesgo y varianza [31].

Ajuste de Hiperparámetros: El ajuste de hiperparámetros es el proceso de seleccionar los valores óptimos para los parámetros de control del modelo que no se aprenden durante el entrenamiento. Técnicas como la búsqueda en cuadrícula (grid search) y la búsqueda aleatoria (random search) son métodos comunes para encontrar los mejores hiperparámetros. Recientemente, se han desarrollado métodos más

avanzados como la optimización bayesiana, que utilizan un enfoque probabilístico para explorar el espacio de hiperparámetros de manera más eficiente [55].

La regularización y la prevención del sobreajuste son componentes esenciales en el desarrollo de modelos de ML robustos y precisos. Aplicar estas técnicas adecuadamente asegura que el modelo no solo funcione bien en los datos de entrenamiento, sino que también sea capaz de generalizar a datos nuevos y desconocidos.

2.7. Infraestructura y Herramientas para Machine Learning

El desarrollo e implementación de modelos de ML requieren una infraestructura robusta y un conjunto de herramientas adecuadas que faciliten cada etapa del ciclo de vida del modelo, desde la recopilación de datos hasta el despliegue y el monitoreo.

Frameworks y Bibliotecas: Existen numerosos frameworks y bibliotecas que simplifican el desarrollo de modelos de ML. Algunos de los más populares incluyen:

- **TensorFlow:** Desarrollado por Google, TensorFlow es un framework de código abierto que permite construir, entrenar y desplegar modelos de ML. Ofrece una amplia gama de herramientas y recursos para el aprendizaje profundo y se integra bien con otras plataformas de Google, como Google Cloud Platform.
- **PyTorch:** Desarrollado por Facebook, PyTorch es conocido por su flexibilidad y facilidad de uso. Es ampliamente utilizado en investigación y producción, y es particularmente popular en la comunidad de aprendizaje profundo debido a su compatibilidad con dinámicas de gráficos computacionales.
- **Scikit-Learn:** Scikit-Learn es una biblioteca de código abierto para ML en Python que proporciona herramientas simples y eficientes para el análisis de datos y modelado predictivo. Es ideal para tareas de clasificación, regresión y clustering, y se integra bien con otras bibliotecas científicas de Python, como NumPy y SciPy.
- **Keras:** Keras es una API de alto nivel para construir y entrenar modelos de redes neuronales. Originalmente desarrollada como una interfaz para Tensor-

Flow, Keras facilita la creación rápida de prototipos y es fácil de usar, lo que la hace ideal para principiantes y expertos.

- **XGBoost:** XGBoost es una biblioteca optimizada para el algoritmo de boosting, que ha demostrado ser extremadamente eficaz en competiciones de ML y en aplicaciones del mundo real. Es conocido por su velocidad y rendimiento, y se utiliza comúnmente en tareas de clasificación y regresión.

Uso de GPUs y TPUs: Los Modelos de Aprendizaje (MA) profundo a menudo requieren una gran cantidad de poder de procesamiento, especialmente durante el entrenamiento. Las Unidades de Procesamiento Gráfico (GPUs) y las Unidades de Procesamiento Tensorial (TPUs) son hardware especializado que acelera significativamente estos cálculos. Frameworks como TensorFlow y PyTorch están diseñados para aprovechar estas unidades, permitiendo el entrenamiento de modelos más grandes y complejos en menos tiempo [56].

Computación en la Nube: La computación en la nube ha transformado el campo del ML al proporcionar acceso bajo demanda a recursos computacionales escalables. Plataformas como Amazon Web Services (AWS), Google Cloud Platform (GCP) y Microsoft Azure ofrecen servicios específicos para ML, incluyendo almacenamiento de datos, herramientas de desarrollo y capacidades de procesamiento de alto rendimiento. Estos servicios permiten a las organizaciones entrenar y desplegar modelos a gran escala sin la necesidad de invertir en infraestructura física costosa [57].

Plataformas de Desarrollo Colaborativo: Herramientas como Jupyter Notebooks y Google Colab han facilitado enormemente el desarrollo y la experimentación en ML. Estas plataformas permiten a los desarrolladores escribir y ejecutar código en fragmentos interactivos, visualizar datos y resultados, y compartir su trabajo con otros. Google Colab, en particular, ofrece acceso gratuito a GPUs, lo que permite a los usuarios entrenar modelos complejos de manera eficiente [58].

Manejo de Datos y Preprocesamiento: La infraestructura de ML también debe abordar el manejo y preprocesamiento de grandes volúmenes de datos. Herramientas como Apache Hadoop y Apache Spark son ampliamente utilizadas para el procesamiento distribuido de grandes conjuntos de datos. Estas herramientas permiten el

procesamiento paralelo de datos a gran escala, lo que es esencial para preparar datos para el entrenamiento de modelos de ML [57].

Gestión del Ciclo de Vida del Modelo: Herramientas como MLflow, Kube-flow y TensorBoard ayudan a gestionar el ciclo de vida completo de los modelos de ML, desde la experimentación y el seguimiento de versiones hasta el despliegue y el monitoreo. MLflow, por ejemplo, ofrece una plataforma para registrar experimentos, empaquetar código en formatos reproducibles y desplegar modelos en diversos entornos. Kubeflow es una plataforma de código abierto diseñada para facilitar el despliegue de flujos de trabajo de ML en Kubernetes, mientras que TensorBoard proporciona herramientas de visualización para el análisis del entrenamiento y la evaluación de modelos de TensorFlow [59].

Automatización de ML (AutoML): AutoML es una tendencia emergente que busca automatizar las partes más tediosas y técnicas del proceso de desarrollo de ML, como la selección de modelos y el ajuste de hiperparámetros. Herramientas como Google Cloud AutoML, AutoKeras y H2O.ai permiten a los usuarios construir modelos de ML de alta calidad con mínima intervención manual, democratizando el acceso al ML avanzado [60].

Monitoreo y Mantenimiento en Producción: Una vez que los modelos de ML se despliegan en producción, es crucial monitorearlos y mantenerlos. Herramientas como Prometheus y Grafana pueden integrarse con los modelos de ML para monitorear su rendimiento y detectar posibles degradaciones. Además, el reentrenamiento periódico de modelos con datos nuevos es una práctica común para asegurar que los modelos sigan siendo precisos y relevantes [61].

La infraestructura y las herramientas adecuadas son fundamentales para el éxito de los proyectos de ML. A medida que el campo continúa evolucionando, la disponibilidad de recursos avanzados y plataformas integradas seguirá facilitando el desarrollo de soluciones innovadoras y eficaces.

2.8. Implicaciones Éticas del Uso de Algoritmos en el Ambiente Educativo

El uso de algoritmos de predicción basados en ML en el ámbito educativo presenta una serie de oportunidades para mejorar la educación mediante la personalización del aprendizaje y la identificación temprana de estudiantes en riesgo [62], [63]. Sin embargo, estas ventajas vienen acompañadas de importantes implicaciones éticas que no pueden pasarse por alto. Es crucial que el desarrollo y la implementación de algoritmos en el contexto educativo se realicen bajo estrictos estándares éticos, ya que estamos tratando con datos sensibles y con decisiones que pueden influir directamente en el futuro académico de los estudiantes [64].

En primer lugar, una de las principales preocupaciones éticas es la **privacidad y protección de los datos personales de los estudiantes**. Los modelos de ML dependen de grandes volúmenes de datos para su entrenamiento, y en el ámbito educativo, estos datos suelen incluir información sensible como calificaciones, género, edad, hábitos de estudio, y en algunos casos, datos socioeconómicos. El manejo inadecuado de esta información puede vulnerar la privacidad de los estudiantes. Para proteger los derechos de los individuos, es fundamental cumplir con normativas como el Reglamento General de Protección de Datos (GDPR) [64], que establece principios estrictos sobre el uso, almacenamiento y manejo de datos personales. Además, las instituciones educativas deben implementar técnicas robustas de anonimización o pseudonimización que aseguren que la identidad de los estudiantes no pueda ser fácilmente inferida a partir de los datos utilizados en los modelos de ML.

Otro aspecto ético relevante es el **sesgo algorítmico**. Los algoritmos de ML se entrenan a partir de datos históricos, y si estos datos contienen sesgos, los modelos tienden a replicarlos o incluso amplificarlos [28]. En el ámbito educativo, esto podría tener consecuencias graves, ya que ciertos grupos de estudiantes, como aquellos provenientes de minorías étnicas o con bajos recursos, podrían recibir predicciones menos favorables debido a sesgos inherentes en los datos. Estos sesgos pueden manifestarse en la asignación de recursos educativos, en el acceso a oportunidades, o en la forma en que se interviene para apoyar a los estudiantes. Para mitigar es-

tos efectos, es esencial aplicar técnicas que permitan detectar y corregir el sesgo en los modelos predictivos. Herramientas como *Equalized Odds* pueden ser útiles para asegurar que las predicciones algorítmicas no favorezcan injustamente a un grupo sobre otro. Asimismo, la implementación de auditorías algorítmicas periódicas puede ayudar a monitorear el comportamiento de los modelos y garantizar que no estén perpetuando desigualdades [65].

La **transparencia y explicabilidad** de los modelos de ML también es una consideración ética fundamental. En el contexto educativo, los docentes y administradores necesitan comprender cómo y por qué los modelos toman ciertas decisiones, especialmente cuando esas decisiones afectan directamente al futuro académico de los estudiantes [66]. Modelos opacos, conocidos como “cajas negras”, pueden generar desconfianza en las herramientas tecnológicas y dificultar su adopción. Para garantizar que los algoritmos de ML sean transparentes, es recomendable utilizar métodos de explicabilidad, como *LIME* (*Local Interpretable Model-agnostic Explanations*) o *SHAP* (*SHapley Additive exPlanations*) [26], que permiten a los usuarios entender las predicciones realizadas por los modelos y los factores que influyeron en dichas predicciones. Además, el GDPR garantiza el derecho a una explicación cuando una decisión automatizada tiene un impacto significativo en el individuo. Por lo tanto, las instituciones educativas que implementen ML deben asegurarse de que los estudiantes y sus familias puedan acceder a explicaciones claras y comprensibles sobre cómo los algoritmos influyen en su proceso educativo.

Otro punto crítico es el **uso ético de las predicciones para la intervención educativa**. Si bien la predicción temprana del rendimiento académico permite identificar a estudiantes en riesgo y proporcionarles apoyo antes de que sea demasiado tarde, también existe el riesgo de que estas predicciones sean utilizadas para etiquetar a los estudiantes o tratarlos de manera diferenciada. Esto puede llevar a la estigmatización de aquellos estudiantes que son clasificados como “de bajo rendimiento”, lo que podría afectar su autoestima o generar un ambiente de discriminación. Es importante que los resultados de los algoritmos de ML se usen de manera cuidadosa, siempre como una herramienta complementaria para los educadores, y no como una base única para tomar decisiones [67]. Las intervenciones deben estar diseñadas

para apoyar a los estudiantes sin crear divisiones o etiquetas negativas.

Finalmente, la **supervisión humana** es indispensable en cualquier sistema automatizado en el ámbito educativo. Las decisiones críticas que afectan el desarrollo académico de los estudiantes no deben ser completamente automatizadas. Aunque los algoritmos de ML pueden proporcionar insights valiosos, es esencial que las decisiones finales sean revisadas y validadas por profesionales de la educación, quienes pueden tener en cuenta factores contextuales que los algoritmos no pueden captar. De esta manera, el uso de ML no desplaza el juicio profesional, sino que lo complementa, permitiendo una intervención educativa más informada y equilibrada [68].

En resumen, la implementación de algoritmos de ML en el ámbito educativo puede ofrecer grandes ventajas, pero también plantea importantes retos éticos. La protección de los datos personales, la corrección de sesgos, la transparencia, el uso adecuado de las predicciones y la supervisión humana son aspectos que deben ser cuidadosamente gestionados para asegurar que la tecnología mejore el proceso educativo sin comprometer los derechos ni el bienestar de los estudiantes. El cumplimiento de estos principios éticos garantizará que las herramientas de ML contribuyan a una educación más justa, equitativa y eficiente.

3. Minería de datos

3.1. Introducción a la Minería de Datos

La MD y el ML son disciplinas complementarias en el campo del análisis de datos, pero difieren en sus enfoques y objetivos. La MD se centra en descubrir patrones y conocimientos a partir de grandes volúmenes de datos utilizando técnicas de estadística, bases de datos y aprendizaje automático. En contraste, el ML se enfoca en desarrollar algoritmos que permitan a las computadoras aprender y hacer predicciones basadas en datos. Mientras que el ML es una herramienta poderosa dentro de la MD, esta última abarca un espectro más amplio de técnicas y aplicaciones.

La MD implica varias etapas, incluyendo la recopilación y preparación de datos, el descubrimiento de patrones, la validación y evaluación de estos patrones, y la presentación del conocimiento extraído. Este proceso estructurado asegura que los resultados obtenidos sean precisos, útiles y aplicables en contextos del mundo real. Por otro lado, el ML se centra en la creación de modelos predictivos que mejoren su desempeño a través de la experiencia y la exposición a nuevos datos.

Uno de los objetivos principales de la MD es identificar relaciones y patrones significativos en los datos que no son evidentes a simple vista. Estos patrones pueden incluir asociaciones, secuencias, clasificaciones, clusters y desviaciones. En comparación, el ML se orienta más hacia la predicción y la toma de decisiones automatizada, utilizando algoritmos como redes neuronales, máquinas de vectores de soporte y árboles de decisión. Mientras que la MD puede emplear ML para mejorar la precisión de sus descubrimientos, también incluye métodos no supervisados y técnicas estadísticas que no necesariamente se centran en la predicción [11].

El proceso de MD sigue un flujo de trabajo estructurado, conocido como el proceso de extracción de conocimiento a partir de bases de datos (KDD, por sus siglas en inglés). Este proceso consta de los siguientes pasos:

1. **Selección de Datos:** Identificación y recolección de datos relevantes para el

análisis.

2. **Preprocesamiento de Datos:** Limpieza y transformación de los datos para eliminar ruido e inconsistencias.
3. **Transformación de Datos:** Conversión de los datos en un formato adecuado para el análisis, incluyendo la reducción de dimensionalidad y la selección de características.
4. **Minería de Datos:** Aplicación de algoritmos para extraer patrones y modelos de los datos.
5. **Evaluación e Interpretación:** Análisis de los patrones descubiertos para determinar su validez y utilidad.
6. **Presentación del Conocimiento:** Presentación de los resultados de manera comprensible y útil para la toma de decisiones [69].

Aunque tanto la MD como el ML se benefician del uso de grandes volúmenes de datos, sus aplicaciones y metodologías pueden diferir significativamente. Por ejemplo, en el marketing, la MD puede utilizarse para identificar segmentos de clientes y desarrollar estrategias de marketing personalizadas, mientras que el ML se emplea para predecir la probabilidad de compra de un cliente individual basándose en su comportamiento pasado [70].

En el ámbito de la atención médica, la MD puede ayudar a descubrir relaciones entre síntomas y enfermedades, mejorando el diagnóstico y tratamiento, mientras que el ML puede ser utilizado para desarrollar modelos predictivos que anticipen la evolución de una enfermedad en base a los datos del paciente. Estas diferencias destacan cómo la MD y el ML pueden trabajar juntos para ofrecer una comprensión más completa y detallada de los datos [25].

En resumen, la MD es una disciplina amplia que abarca una variedad de técnicas para extraer conocimiento útil de grandes conjuntos de datos, incluyendo, pero no limitándose a, métodos de ML. Su capacidad para descubrir patrones ocultos y relaciones en los datos la convierte en una herramienta invaluable en la era de la información, complementando y ampliando las capacidades del ML.

3.2. Técnicas de Minería de Datos

Las técnicas de MD son fundamentales para extraer patrones y conocimiento de grandes volúmenes de datos. Estas técnicas se dividen en varias categorías, cada una adecuada para diferentes tipos de problemas y datos. A continuación, se describen algunas de las técnicas más comunes utilizadas en la MD.

Clasificación: La clasificación es una técnica supervisada utilizada para predecir la etiqueta de una instancia basándose en características observables. Los algoritmos de clasificación se entrenan utilizando un dataset etiquetados y aprenden a asignar nuevas instancias a una de las clases predefinidas. Algunos de los algoritmos de clasificación más populares incluyen los árboles de decisión, las máquinas de vectores de soporte (SVM) y las redes neuronales [71].

Regresión: La regresión es una técnica supervisada que se utiliza para predecir un valor continuo en lugar de una clase discreta. La regresión lineal es uno de los métodos más simples y comunes, pero también existen técnicas más avanzadas como la regresión polinómica y la regresión de bosques aleatorios. La regresión se utiliza en diversas aplicaciones, desde la predicción de precios hasta la estimación de tendencias de mercado [36].

Agrupamiento (Clustering): El agrupamiento es una técnica no supervisada que agrupa instancias similares en clusters. A diferencia de la clasificación, el agrupamiento no requiere etiquetas de clase en los datos de entrenamiento. Algoritmos como K-means, DBSCAN y los algoritmos jerárquicos son ampliamente utilizados en tareas de agrupamiento. Esta técnica es útil para segmentar mercados, detectar patrones en datos de clientes y analizar comportamientos en redes sociales [72].

Asociación: La asociación es una técnica utilizada para descubrir relaciones interesantes entre variables en grandes bases de datos. Las reglas de asociación se utilizan comúnmente en el análisis de la cesta de la compra para identificar productos que se compran juntos con frecuencia. El algoritmo Apriori es uno de los más conocidos para la generación de reglas de asociación. Esta técnica es valiosa para la recomendación de productos y el análisis de hábitos de consumo [73].

Detección de Anomalías: La detección de anomalías, o detección de outliers, es una técnica utilizada para identificar instancias que no se ajustan a un patrón esperado. Estas anomalías pueden indicar fraudes, fallos del sistema o comportamientos inusuales. Los métodos para la detección de anomalías incluyen técnicas estadísticas, algoritmos basados en distancias y MA automático. La detección de anomalías es crucial en la seguridad informática, la detección de fraudes y el mantenimiento predictivo [74].

Reducción de Dimensionalidad: La reducción de dimensionalidad es una técnica utilizada para reducir el número de variables en un dataset mientras se preserva la mayor cantidad de información posible. Esto es importante para simplificar modelos, reducir el costo computacional y eliminar ruido. El Análisis de Componentes Principales (PCA) y el Análisis Discriminante Lineal (LDA) son métodos comunes de reducción de dimensionalidad. Estas técnicas se aplican en la visualización de datos, la compresión de datos y la preprocesamiento de datos para otros algoritmos de MD [75].

Minería de Textos: La minería de textos implica la extracción de información y patrones a partir de datos textuales. Esta técnica es especialmente relevante en la era digital, donde una gran cantidad de información se almacena en formato de texto. Las técnicas de minería de textos incluyen la extracción de entidades, el análisis de sentimientos y la clasificación de documentos. Los algoritmos de procesamiento del lenguaje natural (NLP) y el aprendizaje profundo son herramientas clave en la minería de textos [76].

Cada una de estas técnicas de MD ofrece ventajas específicas y es adecuada para diferentes tipos de problemas y contextos. La selección de la técnica correcta depende de la naturaleza de los datos y de los objetivos del análisis. A medida que la tecnología avanza, se desarrollan nuevas técnicas y algoritmos que amplían las capacidades de la MD, permitiendo un análisis más profundo y preciso de datos complejos y diversos.

3.3. Herramientas y Plataformas para Minería de Datos

La MD requiere herramientas y plataformas robustas que faciliten la extracción, procesamiento y análisis de grandes volúmenes de datos. A continuación, se describen algunas de las herramientas y plataformas más populares utilizadas en la MD.

Python y R: Python y R son lenguajes de programación ampliamente utilizados en la MD y la ciencia de datos. Python es conocido por su simplicidad y la riqueza de sus bibliotecas, como Pandas para la manipulación de datos, Scikit-Learn para ML y Matplotlib para la visualización. R es un lenguaje estadístico que ofrece una vasta colección de paquetes para el análisis de datos y la visualización, como ggplot2 y dplyr. Ambos lenguajes son populares entre investigadores y profesionales debido a su flexibilidad y extensibilidad.

Microsoft Azure ML: Microsoft Azure Machine Learning es un servicio en la nube que permite a los usuarios construir, entrenar y desplegar modelos de ML a escala. Proporciona herramientas para el preprocesamiento de datos, la selección de características y la evaluación de modelos, y soporta una variedad de algoritmos y frameworks de ML. Azure Machine Learning es ideal para empresas que buscan una solución escalable y gestionada para sus necesidades de análisis de datos.

Weka: Weka (Waikato Environment for Knowledge Analysis) es una suite de software de MD desarrollada en la Universidad de Waikato, Nueva Zelanda. Proporciona una colección de algoritmos de ML y herramientas de visualización que pueden ser aplicadas directamente a conjuntos de datos. Weka es ampliamente utilizado en la investigación y en la enseñanza debido a su interfaz gráfica de usuario intuitiva y su extensa documentación.

RapidMiner: RapidMiner es una plataforma de ciencia de datos que ofrece un entorno integrado para la preparación de datos, ML y despliegue de modelos. Proporciona una interfaz gráfica de usuario que permite a los usuarios construir flujos de trabajo de análisis de datos sin necesidad de programar. RapidMiner es utilizado tanto por investigadores como por profesionales de la industria debido a su flexibilidad y capacidad para manejar grandes volúmenes de datos.

3.4. Aplicaciones de la Minería de Datos

La MD tiene una amplia gama de aplicaciones en diversos sectores, ayudando a las organizaciones a extraer valor de sus datos y a tomar decisiones informadas. A continuación, se presentan algunas de las aplicaciones más comunes de la MD.

Marketing y Ventas: En marketing y ventas, la MD se utiliza para analizar el comportamiento de los clientes, segmentar mercados y desarrollar campañas de marketing personalizadas. Los modelos predictivos pueden identificar clientes potenciales, predecir el abandono de clientes y recomendar productos basados en patrones de compra. Las técnicas de asociación, como el análisis de la cesta de la compra, ayudan a descubrir productos que se compran juntos con frecuencia, permitiendo a los minoristas optimizar la disposición de productos y las promociones [77].

Detección de Fraudes: La MD es crucial en la detección de fraudes en sectores como la banca y los seguros. Los algoritmos de detección de anomalías pueden identificar transacciones sospechosas y comportamientos inusuales que pueden indicar actividades fraudulentas. Los modelos de clasificación y las redes neuronales se utilizan para analizar patrones históricos de fraude y mejorar la precisión de la detección. Estas técnicas ayudan a las organizaciones a reducir pérdidas y a proteger a sus clientes [78].

Atención Médica: En el campo de la atención médica, la MD se aplica para mejorar la calidad del cuidado de los pacientes y optimizar las operaciones hospitalarias. Los modelos predictivos pueden ayudar en el diagnóstico de enfermedades, la identificación de factores de riesgo y la personalización de tratamientos. La MD también se utiliza para analizar grandes volúmenes de datos clínicos y descubrir relaciones entre síntomas, diagnósticos y resultados de tratamientos. Esto facilita la toma de decisiones basada en evidencia y la investigación médica [79].

Finanzas y Banca: En el sector financiero, la MD se emplea para gestionar riesgos, predecir tendencias del mercado y optimizar carteras de inversión. Los modelos de regresión y las series temporales se utilizan para predecir precios de acciones, tasas de interés y otros indicadores financieros. La MD también ayuda a las instituciones

financieras a personalizar ofertas de productos y a mejorar la satisfacción del cliente mediante el análisis de patrones de comportamiento financiero [80].

Telecomunicaciones: En la industria de las telecomunicaciones, la MD se utiliza para gestionar redes, mejorar el servicio al cliente y desarrollar nuevos productos. Los modelos predictivos pueden identificar clientes que están en riesgo de abandonar el servicio y sugerir acciones para retenerlos. La MD también se aplica en el análisis de fallos de red y en la optimización de recursos de red para mejorar la calidad del servicio [81].

Educación: En el ámbito educativo, la MD se utiliza para mejorar los procesos de enseñanza y aprendizaje. Los sistemas de recomendación pueden sugerir recursos educativos personalizados a los estudiantes, y los modelos predictivos pueden identificar estudiantes en riesgo de bajo rendimiento o abandono. La MD también se emplea para analizar datos de encuestas y evaluaciones, proporcionando información valiosa para la mejora de los currículos y las estrategias pedagógicas [82].

E-commerce: En el comercio electrónico, la MD se aplica para personalizar la experiencia del usuario, mejorar la satisfacción del cliente y aumentar las ventas. Los sistemas de recomendación, basados en técnicas de filtrado colaborativo y análisis de preferencias, sugieren productos relevantes a los clientes. La MD también ayuda a optimizar los motores de búsqueda y a analizar el comportamiento de los usuarios en el sitio web, lo que permite a las empresas mejorar la usabilidad y la efectividad de sus plataformas [83].

Seguridad Nacional: La MD juega un papel importante en la seguridad nacional, ayudando a las agencias gubernamentales a detectar amenazas y prevenir actos delictivos. Los algoritmos de análisis de redes sociales pueden identificar patrones de comportamiento sospechosos y conexiones entre individuos. La MD también se utiliza para analizar grandes volúmenes de datos de inteligencia y detectar actividades potencialmente peligrosas, mejorando la capacidad de respuesta ante amenazas [84].

Estas aplicaciones demuestran la versatilidad y el impacto de la MD en diferentes sectores. A medida que las organizaciones continúan acumulando grandes volúmenes de datos, la capacidad de extraer conocimiento valioso a partir de estos datos se

vuelve cada vez más crucial. La MD ofrece herramientas y técnicas poderosas para transformar datos en información útil, mejorando la toma de decisiones y la eficiencia operativa.

3.5. Desafíos y Tendencias Futuras en Minería de Datos

La MD enfrenta varios desafíos que deben ser abordados para maximizar su efectividad y aplicabilidad. Además, el campo está en constante evolución, con nuevas tendencias que prometen transformar la forma en que se analizan y utilizan los datos.

Desafíos en Minería de Datos:

Calidad de los Datos: La calidad de los datos es un factor crítico en el éxito de cualquier proyecto de MD. Los datos incompletos, ruidosos y sesgados pueden llevar a resultados incorrectos o engañosos. El preprocesamiento de datos, que incluye la limpieza, integración y transformación de datos, es esencial para mejorar la calidad de los datos y garantizar la precisión de los modelos de MD.

Escalabilidad: A medida que las organizaciones acumulan grandes volúmenes de datos, la escalabilidad de las técnicas de MD se convierte en un desafío significativo. Las herramientas y algoritmos deben ser capaces de manejar eficientemente conjuntos de datos masivos sin comprometer el rendimiento. Las soluciones incluyen el uso de plataformas de procesamiento distribuido como Apache Hadoop y Apache Spark, y el desarrollo de algoritmos de MD escalables.

Privacidad y Seguridad: La privacidad y seguridad de los datos son preocupaciones importantes, especialmente cuando se trata de datos sensibles como información personal y financiera. Es esencial implementar medidas de seguridad robustas para proteger los datos contra accesos no autorizados y ataques cibernéticos.

Interpretabilidad: La interpretabilidad de los modelos de MD es crucial para generar confianza en las decisiones basadas en datos. Los modelos complejos, como las redes neuronales profundas, a menudo se consideran cajas negras debido a su falta de transparencia. Se están desarrollando técnicas para mejorar la interpretabilidad de

estos modelos, como los métodos de explicabilidad de modelos y las visualizaciones de datos intuitivas.

Integración de Datos Heterogéneos: En muchos casos, los datos provienen de múltiples fuentes con diferentes formatos y estructuras. La integración de estos datos heterogéneos es un desafío, ya que requiere técnicas avanzadas de fusión y transformación de datos para crear un dataset coherente y utilizable. Las soluciones incluyen el uso de ontologías y modelos de datos unificados para facilitar la integración de datos de diferentes fuentes.



4. Inteligencia Artificial

La IA ha emergido como una de las áreas más dinámicas y transformadoras en el campo de la informática. Desde sus inicios, la IA ha evolucionado para convertirse en una tecnología omnipresente que influye en numerosos aspectos de la vida moderna. Su capacidad para emular procesos cognitivos humanos ha permitido su aplicación en diversas industrias, desde la medicina hasta la educación, pasando por la industria manufacturera y el entretenimiento.

4.1. Introducción a la Inteligencia Artificial

La IA es un campo de la informática que se ocupa del desarrollo de sistemas capaces de realizar tareas que normalmente requerirían inteligencia humana. Estas tareas incluyen el reconocimiento de patrones, el aprendizaje, la toma de decisiones y la adaptación a nuevas situaciones. El objetivo principal de la IA es emular las capacidades cognitivas humanas en máquinas, permitiendo a las computadoras pensar, razonar y aprender de manera autónoma.

El concepto de IA ha evolucionado significativamente desde sus orígenes a mediados del siglo XX. Uno de los primeros hitos en la historia de la IA fue el trabajo de Alan Turing, quien en 1950 planteó la posibilidad de que las máquinas pudieran realizar tareas cognitivas similares a las de los humanos. Su artículo "Computing Machinery and Intelligence" introdujo el famoso Test de Turing, una prueba diseñada para evaluar si una máquina puede exhibir un comportamiento inteligente comparable al de un ser humano [85]. Este fue uno de los primeros intentos de definir y abordar la inteligencia desde una perspectiva computacional.

A lo largo de las décadas, la IA ha pasado por varios ciclos de entusiasmo y decepción. En los primeros años, los investigadores eran optimistas acerca de la rapidez con la que se podrían lograr avances significativos. Sin embargo, las limitaciones tecnológicas y la complejidad inherente a la creación de sistemas verdaderamente inteligentes resultaron ser desafíos más formidables de lo que se había anticipado.

Estos periodos de estancamiento, conocidos como los "inviernos de la IA", llevaron a una disminución en el financiamiento y el interés en el campo [86]. Sin embargo, a medida que las tecnologías han avanzado, la IA ha resurgido con una fuerza renovada, impulsada por el crecimiento exponencial en la capacidad de procesamiento de datos, el desarrollo de algoritmos más sofisticados y la disponibilidad de grandes volúmenes de datos (Big Data) [87].

Hoy en día, la IA se ha convertido en una parte integral de muchas industrias, desempeñando un papel clave en la automatización de procesos, la optimización de recursos y la personalización de servicios. Desde la salud hasta las finanzas, pasando por la educación y la seguridad, las aplicaciones de la IA son vastas y variadas. Una de las áreas más destacadas es el ML o aprendizaje automático, que permite a las máquinas aprender y mejorar su desempeño a partir de la experiencia sin necesidad de ser programadas explícitamente para cada tarea. El aprendizaje automático se basa en algoritmos que pueden identificar patrones en grandes conjuntos de datos y ajustar sus procesos internos para mejorar la precisión de sus predicciones o decisiones [25].

La IA se divide generalmente en dos categorías principales: IA débil e IA fuerte. La IA débil, también conocida como IA narrow, se refiere a sistemas diseñados para realizar tareas específicas de manera eficiente, pero que no poseen la capacidad de realizar tareas fuera de su dominio especializado. Un ejemplo clásico de IA débil es el asistente virtual en un teléfono móvil, que puede realizar tareas como responder preguntas, establecer recordatorios y enviar mensajes, pero no puede aprender o adaptarse a nuevos contextos de manera significativa fuera de las funciones para las que fue programado [88].

Por otro lado, la IA fuerte es un concepto que describe sistemas con capacidades cognitivas equiparables a las de un ser humano. Estos sistemas no solo podrían realizar una amplia variedad de tareas, sino que también podrían aprender y adaptarse de manera flexible a diferentes contextos. Aunque la IA fuerte sigue siendo más un objetivo teórico que una realidad práctica, su desarrollo implicaría un cambio fundamental en cómo interactuamos con las máquinas y las tareas que podemos delegarles [89].

Además de la IA débil y fuerte, también se discute la posibilidad de la superinteligencia, que se refiere a una IA que supera ampliamente la inteligencia humana en todos los aspectos, incluyendo la creatividad, la toma de decisiones y la resolución de problemas. Aunque la superinteligencia es un concepto teórico, ha generado un considerable debate ético y filosófico, dado su potencial para transformar la sociedad de maneras que aún no podemos prever [89].

El continuo desarrollo de la IA plantea una serie de preguntas y desafíos, tanto técnicos como éticos. A medida que las máquinas se vuelven más capaces y autónomas, surge la necesidad de considerar cómo se utilizará esta tecnología y cómo se regulará para garantizar que beneficie a la sociedad en su conjunto. La privacidad, la seguridad, el sesgo en los algoritmos y el impacto en el empleo son solo algunas de las cuestiones que deben ser abordadas a medida que la IA continúa evolucionando [90].

4.2. Historia de la Inteligencia Artificial

La historia de la IA es rica y compleja, llena de descubrimientos, avances tecnológicos y, en ocasiones, periodos de estancamiento conocidos como "inviernos de la IA". La IA tiene sus raíces en la antigua Grecia, donde filósofos como Aristóteles sentaron las bases del razonamiento lógico que eventualmente inspiraría a los primeros científicos computacionales. Sin embargo, fue en el siglo XX cuando la IA comenzó a tomar forma como un campo de estudio distintivo.

El nacimiento oficial de la IA como disciplina se atribuye a la conferencia de Dartmouth en 1956, organizada por John McCarthy, Marvin Minsky, Nathaniel Rochester y Claude Shannon. Este evento fue clave para definir el campo de la IA y marcar el inicio de una era de exploración y desarrollo intensivo en este ámbito [91]. Durante esta conferencia, los asistentes discutieron sobre la posibilidad de crear máquinas que pudieran simular todos los aspectos de la inteligencia humana, sentando las bases para décadas de investigación futura.

En los años que siguieron, la IA experimentó un período de gran optimismo. Du-

rante la década de 1960, se desarrollaron los primeros programas de IA capaces de resolver problemas matemáticos complejos, jugar al ajedrez y manipular símbolos en lenguajes formales. Un ejemplo notable de estos primeros sistemas es el programa Logic Theorist, creado por Allen Newell y Herbert A. Simon, que fue capaz de demostrar teoremas matemáticos al igual que un humano [92].

El entusiasmo inicial por la IA condujo a una oleada de investigaciones y financiamiento en las décadas de 1960 y 1970. Se desarrollaron varias técnicas fundamentales, como los sistemas expertos, que permitían a las computadoras tomar decisiones basadas en reglas predefinidas. Uno de los sistemas expertos más conocidos fue Dendral, desarrollado en la Universidad de Stanford, que ayudó a los químicos a identificar estructuras moleculares utilizando reglas derivadas del conocimiento experto en el campo [88].

A pesar de estos avances, la IA pronto se encontró con serias limitaciones. Los sistemas basados en reglas demostraron ser inflexibles y no podían manejar la ambigüedad o la incertidumbre inherente a muchos problemas del mundo real. Como resultado, la promesa de crear máquinas inteligentes que pudieran igualar o superar la inteligencia humana comenzó a parecer inalcanzable. Esto llevó a un declive en el interés y la financiación durante los años 1970 y 1980, un periodo que se conoce como el primer "invierno de la IA" [86].

A mediados de la década de 1980, la IA experimentó un resurgimiento gracias al éxito de las redes neuronales artificiales, una técnica inspirada en el funcionamiento del cerebro humano. Estas redes, que consisten en capas de neuronas artificiales interconectadas, demostraron ser capaces de aprender patrones complejos a partir de datos. Sin embargo, a pesar del entusiasmo inicial, las redes neuronales enfrentaron dificultades debido a las limitaciones computacionales de la época y a la falta de datos suficientes para entrenar modelos eficaces. Este periodo de declive en el interés llevó al segundo "invierno de la IA" a finales de los años 1980 y principios de los 1990 [87].

El verdadero renacimiento de la IA comenzó a principios del siglo XXI, impulsado por varios factores clave. Primero, el aumento exponencial en la capacidad de

procesamiento de las computadoras permitió a los investigadores crear modelos más complejos y entrenarlos en grandes conjuntos de datos. Segundo, la proliferación de datos disponibles, en gran parte gracias a Internet y a los dispositivos conectados, proporcionó el combustible necesario para que los algoritmos de ML pudieran aprender y generalizar de manera efectiva [19]. Finalmente, la aparición de nuevas técnicas, como el ML y el Análisis Predictivo (AP), transformó la IA en un campo dinámico y en rápido crecimiento [25].

En la actualidad, la IA ha pasado de ser una curiosidad académica a una tecnología central en muchas industrias. Las redes neuronales profundas, una extensión de las redes neuronales tradicionales, han permitido avances significativos en áreas como el PLD, la visión por computadora y los sistemas de recomendación. Empresas de tecnología líderes, como Google, Amazon, y Facebook, han adoptado la IA como una parte fundamental de sus operaciones, utilizando algoritmos avanzados para personalizar experiencias de usuario, optimizar procesos y desarrollar nuevos productos [89].

Además, la IA ha comenzado a desempeñar un papel crucial en campos como la medicina, donde se utiliza para el análisis de imágenes médicas, la predicción de enfermedades y la personalización de tratamientos. En la industria financiera, la IA se emplea para el análisis de riesgos, la detección de fraudes y la optimización de carteras de inversión. Incluso en el ámbito creativo, la IA está comenzando a ser utilizada para generar música, arte y literatura, desafiando nuestras concepciones tradicionales de la creatividad [93].

A pesar de estos avances, la historia de la IA también está marcada por desafíos persistentes. El desarrollo de sistemas de IA que sean verdaderamente autónomos y capaces de razonar a un nivel similar al humano sigue siendo un objetivo elusivo. Además, la rápida adopción de la IA plantea importantes cuestiones éticas y sociales, como el impacto en el empleo, la privacidad de los datos y el sesgo en los algoritmos [90].

4.3. Modelos de Inteligencia Artificial

La IA abarca una amplia variedad de modelos y técnicas que se utilizan para simular la inteligencia humana en máquinas. Estos modelos se diferencian entre sí en función de su complejidad, sus aplicaciones y las formas en que procesan la información. A continuación, se describen algunos de los tipos más relevantes de modelos de IA, que han sido fundamentales en el desarrollo de esta disciplina.

Uno de los modelos más básicos y extendidos en la IA es el **aprendizaje supervisado**. En este enfoque, los modelos se entrenan con un dataset etiquetado, lo que significa que cada entrada de datos va acompañada de una salida deseada conocida. El objetivo del modelo es aprender a mapear las entradas a las salidas correctas, de manera que pueda predecir las salidas correspondientes para nuevas entradas. El aprendizaje supervisado es ampliamente utilizado en tareas como la clasificación de correos electrónicos (spam o no spam), el reconocimiento de voz y la predicción de precios de activos financieros [25].

Por otro lado, el **aprendizaje no supervisado** se ocupa de modelos que trabajan con datos no etiquetados. A diferencia del aprendizaje supervisado, aquí el objetivo es descubrir patrones o estructuras ocultas en los datos sin tener una salida específica como referencia. Los algoritmos de clustering, como el K-means, son un ejemplo común de aprendizaje no supervisado, donde el modelo agrupa los datos en categorías basadas en similitudes inherentes [87]. Este tipo de aprendizaje es útil en aplicaciones como la segmentación de clientes en marketing y la detección de anomalías en la seguridad informática.

El **aprendizaje por refuerzo** es otro tipo de modelo de IA que se basa en la idea de que un agente puede aprender a tomar decisiones óptimas interactuando con su entorno. En lugar de aprender de un dataset fijo, el agente recibe recompensas o penalizaciones en función de las acciones que realiza, lo que le permite mejorar su política de toma de decisiones a lo largo del tiempo. El aprendizaje por refuerzo ha sido clave en el desarrollo de sistemas autónomos, como los robots y los vehículos autoconducidos, donde es crucial aprender comportamientos óptimos en un entorno

dinámico [94].

Las **redes neuronales artificiales** representan un tipo de modelo más complejo inspirado en la estructura del cerebro humano. Estas redes están formadas por capas de neuronas artificiales interconectadas, donde cada neurona recibe entradas, realiza cálculos y transmite una salida a la siguiente capa. Las redes neuronales han demostrado ser extremadamente eficaces en el reconocimiento de patrones complejos y en tareas como la clasificación de imágenes, el procesamiento del lenguaje natural y el reconocimiento de voz [19]. En particular, las redes neuronales profundas, que consisten en múltiples capas, han permitido avances significativos en la IA moderna.

Un modelo de IA particularmente destacado en la última década es el **aprendizaje profundo** (*Deep Learning*), una subcategoría de las redes neuronales artificiales. El aprendizaje profundo ha revolucionado áreas como la visión por computadora, el PLD y los sistemas de recomendación, al permitir que los modelos automaticen la extracción de características a partir de los datos brutos [25]. Los modelos de aprendizaje profundo son capaces de aprender representaciones jerárquicas de los datos, lo que les permite capturar relaciones complejas y mejorar significativamente el rendimiento en tareas desafiantes.

Otro modelo fundamental en la IA es el **árbol de decisión**. Un árbol de decisión es una estructura jerárquica en la que cada nodo interno representa una "pregunta" sobre un atributo de los datos, y cada rama representa la respuesta a esa pregunta, que conduce a los nodos hijos. Las hojas del árbol representan las decisiones finales o clasificaciones. Los árboles de decisión son populares debido a su simplicidad y facilidad de interpretación. Son ampliamente utilizados en clasificación y regresión, y sirven como base para métodos más avanzados como los bosques aleatorios y los métodos de ensamble [87].

Además de estos modelos, existen otros enfoques como los **sistemas expertos**, que se basan en reglas predefinidas y conocimiento experto codificado para tomar decisiones en dominios específicos. Aunque los sistemas expertos han sido superados en muchas aplicaciones por modelos más flexibles como las redes neuronales, siguen siendo útiles en áreas donde el conocimiento experto es difícil de formalizar

de manera algorítmica [88].

Finalmente, un área emergente en la IA es la **IA simbólica**, que se basa en la manipulación de símbolos y reglas lógicas para realizar tareas de razonamiento y resolución de problemas. Aunque este enfoque fue prominente en los primeros días de la IA, ha sido redescubierto recientemente como complemento a las técnicas basadas en datos, especialmente en aplicaciones que requieren interpretabilidad y razonamiento lógico [87].

4.4. Aplicaciones Actuales de la Inteligencia Artificial

La IA ha dejado de ser un concepto teórico para convertirse en una realidad que impacta numerosas industrias y aspectos de la vida cotidiana. Desde la automatización de procesos hasta la personalización de servicios, las aplicaciones actuales de la IA son vastas y variadas. A continuación, se destacan algunas de las áreas donde la IA está teniendo un impacto significativo.

Una de las aplicaciones más notables de la IA se encuentra en el **sector de la salud**. Aquí, la IA se utiliza para el análisis de grandes volúmenes de datos médicos, lo que permite a los profesionales de la salud realizar diagnósticos más precisos y rápidos. Por ejemplo, los algoritmos de ML son capaces de analizar imágenes médicas, como radiografías y resonancias magnéticas, para detectar anomalías que podrían ser indicativas de enfermedades como el cáncer [19]. Además, la IA se está utilizando para personalizar los tratamientos en función del perfil genético de cada paciente, lo que marca el comienzo de la medicina de precisión.

En el **ámbito financiero**, la IA ha transformado la forma en que las instituciones financieras operan. Los algoritmos de AP se emplean para prever movimientos en los mercados financieros, permitiendo a los inversores tomar decisiones informadas y en tiempo real. Asimismo, la IA se utiliza para la detección de fraudes al identificar patrones inusuales en las transacciones financieras que podrían indicar actividades fraudulentas [87]. Esto ha mejorado considerablemente la seguridad y la eficiencia en el manejo de finanzas.

La **industria del comercio electrónico** también se ha beneficiado enormemente de la IA, particularmente en la personalización de la experiencia del usuario. Plataformas como Amazon y Netflix utilizan algoritmos de ML para recomendar productos y contenidos a sus usuarios, basándose en su historial de navegación y compra. Esto no solo mejora la satisfacción del cliente, sino que también aumenta las ventas al dirigir a los usuarios hacia productos que son más propensos a comprar [25]. Además, la IA se emplea en la gestión de inventarios y la optimización de cadenas de suministro, lo que permite a las empresas responder de manera más eficiente a la demanda del mercado.

Otra área clave donde la IA está haciendo una diferencia es en la **seguridad y vigilancia**. Los sistemas de reconocimiento facial y análisis de video impulsados por IA se utilizan para monitorear espacios públicos y privados, mejorando la capacidad de las autoridades para prevenir y responder a incidentes de seguridad. Estos sistemas pueden identificar personas de interés en grandes multitudes o detectar comportamientos anómalos que podrían indicar una amenaza inminente [88]. Sin embargo, estas aplicaciones también han generado un debate sobre la privacidad y los derechos civiles, lo que subraya la necesidad de regulaciones claras en el uso de la IA en la vigilancia.

En el **sector de la educación**, la IA está revolucionando la forma en que los estudiantes aprenden y los profesores enseñan. Las plataformas educativas basadas en IA pueden personalizar el contenido de aprendizaje para cada estudiante, adaptándose a su ritmo y estilo de aprendizaje. Esto permite una experiencia de aprendizaje más eficiente y efectiva. Además, los sistemas de evaluación automatizados pueden proporcionar retroalimentación inmediata a los estudiantes, ayudándolos a mejorar sus habilidades y conocimientos de manera más rápida [94].

Por último, la IA está comenzando a jugar un papel importante en la **industria creativa**. Los algoritmos de IA se utilizan para generar música, arte y literatura, abriendo nuevas posibilidades en la creación artística. Herramientas como DALL-E y GPT, desarrolladas por OpenAI, han demostrado que la IA puede crear imágenes y textos que son indistinguibles de los creados por humanos. Aunque estas tecnologías todavía están en sus primeras etapas, representan un cambio significativo en la forma

en que entendemos la creatividad y el arte.

4.5. Problemas Legales y Morales del Uso de la Inteligencia Artificial

El avance de la IA ha traído consigo una serie de desafíos legales y morales que requieren atención inmediata. A medida que esta tecnología se integra más profundamente en nuestras vidas y en diversas industrias, se hace imprescindible abordar las implicaciones éticas, sociales y legales que conlleva. A continuación, se discuten algunos de los problemas más significativos relacionados con el uso de la IA.

Uno de los problemas más críticos en la IA es el sesgo y la discriminación que pueden surgir de los algoritmos. Los sistemas de IA se entrenan en grandes volúmenes de datos, los cuales a menudo reflejan los prejuicios y desigualdades presentes en la sociedad. Como resultado, los modelos de IA pueden perpetuar e incluso amplificar estos sesgos, conduciendo a decisiones que favorecen a ciertos grupos sobre otros. Este tipo de sesgo ha sido documentado en diversos contextos, como en algoritmos de contratación que priorizan a candidatos masculinos sobre femeninos debido a sesgos presentes en los datos históricos de contratación [95]. La cuestión legal aquí se centra en cómo regular y mitigar estos sesgos para asegurar que los sistemas de IA sean justos y equitativos. Aunque algunas regiones, como la Unión Europea, han comenzado a desarrollar marcos regulatorios específicos, la implementación de estas regulaciones presenta desafíos significativos, especialmente en términos de supervisión y cumplimiento [96].

La privacidad y la seguridad de los datos representan otro aspecto fundamental en la discusión sobre la IA. Los modelos de IA dependen de vastas cantidades de datos, muchos de los cuales son de naturaleza personal y sensible. Esto plantea serias preocupaciones sobre la protección de la privacidad de los individuos y el potencial uso indebido de la información recopilada. La recopilación y el procesamiento masivo de datos pueden llevar a la creación de perfiles detallados de las personas, que podrían ser utilizados para la vigilancia estatal o la explotación comercial sin el consentimiento adecuado. El Reglamento General de Protección de Datos (GDPR)

en Europa es un ejemplo de cómo los legisladores están tratando de abordar estos problemas, estableciendo estrictas normas para la protección de datos personales [97]. Sin embargo, a medida que la IA sigue evolucionando, las leyes existentes pueden necesitar adaptaciones para mantener el ritmo con los avances tecnológicos y proteger adecuadamente los derechos de los individuos.

La transparencia y la explicabilidad de los algoritmos son otro desafío importante en el uso de la IA. Muchos modelos de IA, especialmente aquellos basados en aprendizaje profundo, operan como “cajas negras”, lo que significa que incluso los desarrolladores pueden tener dificultades para entender cómo y por qué se toma una decisión específica. Esta falta de explicabilidad es especialmente problemática en áreas donde las decisiones tienen consecuencias graves, como en la justicia penal o la atención médica. Desde una perspectiva legal y ética, es crucial que los sistemas de IA sean transparentes y que las decisiones puedan ser explicadas y justificadas. La falta de transparencia puede socavar la confianza en la IA y dificultar su adopción por parte de instituciones y usuarios. Propuestas legales y técnicas están emergiendo para desarrollar modelos de IA que sean más interpretables y que permitan una mejor comprensión de sus procesos internos [98].

La responsabilidad legal es otro tema complejo cuando se trata de IA. Si un sistema de IA comete un error o toma una decisión perjudicial, surge la cuestión de quién debe ser considerado responsable: ¿el desarrollador del algoritmo, la empresa que lo implementa, o el sistema en sí? Esta es una pregunta particularmente delicada en el contexto de los vehículos autónomos, donde los errores pueden tener consecuencias graves. La responsabilidad en la IA debe ser compartida entre los desarrolladores, las empresas y los reguladores que deben garantizar que los sistemas se implementen de manera segura y ética. Sin embargo, las leyes actuales no están completamente equipadas para manejar estas complejidades, lo que sugiere la necesidad de desarrollar marcos legales específicos para abordar estas cuestiones [99].

Finalmente, las implicaciones éticas y sociales del uso generalizado de la IA son profundas. La automatización impulsada por la IA puede llevar a la pérdida de empleos en ciertos sectores, exacerbando la desigualdad económica y creando nuevas divisiones sociales. Además, la capacidad de la IA para influir en las decisiones hu-

manas plantea preocupaciones sobre la manipulación y la erosión del libre albedrío. Es necesario abordar estas preocupaciones para asegurar que la IA beneficie a la sociedad en su conjunto y no solo a una élite tecnológica o económica. Esto podría incluir la promoción de la educación y el reentrenamiento de la fuerza laboral para adaptarse a un mundo impulsado por la IA, así como la implementación de políticas que aseguren una distribución equitativa de los beneficios económicos de la IA [100].

5. Learning analytics

5.1. Introducción y Fundamentos Teóricos de Learning Analytics

LA es una disciplina emergente dedicada a la medición, recopilación, análisis y reporte de datos sobre estudiantes y sus contextos de aprendizaje. Su objetivo es comprender y optimizar tanto el proceso de aprendizaje como los entornos educativos en los que ocurre.

El propósito principal de LA es utilizar datos para mejorar la enseñanza y el aprendizaje. A través del análisis de datos educativos, se pueden identificar patrones y tendencias que permiten a los educadores tomar decisiones informadas y personalizar la instrucción según las necesidades individuales de los estudiantes [101]. Este enfoque no solo enriquece la experiencia educativa, sino que también promueve la eficiencia en la administración de recursos educativos.

Historia y Evolución de Learning Analytics: El concepto de LA ha evolucionado significativamente desde principios de la década de 2000. Inicialmente, se enfocaba en el análisis de datos de aprendizaje en línea y en la evaluación del desempeño de los estudiantes. Con el tiempo, la disponibilidad de grandes volúmenes de datos educativos y el avance de las tecnologías de análisis de datos han ampliado el alcance de LA, permitiendo aplicaciones más sofisticadas y personalizadas [102]. Hoy en día, LA se aplica en diversas áreas educativas, desde la educación primaria hasta la formación profesional y el aprendizaje a lo largo de la vida.

Teorías y Modelos Subyacentes: LA se basa en diversas teorías y modelos que guían su implementación y aplicación. Entre estas teorías se encuentran:

- **Constructivismo:** Esta teoría del aprendizaje sugiere que los estudiantes construyen activamente su conocimiento a través de la interacción con el entorno y la reflexión sobre sus experiencias. LA puede apoyar el constructivismo proporcionando retroalimentación inmediata y recursos personalizados que faciliten la construcción del conocimiento [103].
- **Conectivismo:** Propuesta por Siemens, esta teoría del aprendizaje para la era digital enfatiza la importancia de las redes y conexiones en el proceso de aprendizaje. LA facilita el conectivismo analizando las interacciones de los estudiantes en entornos de aprendizaje en línea e identificando redes de colaboración efectivas [104].
- **Teoría del Aprendizaje Autorregulado:** Esta teoría destaca la importancia de que los estudiantes sean conscientes de su proceso de aprendizaje y tomen control activo de sus estrategias de aprendizaje. LA puede apoyar la autorregulación proporcionando a los estudiantes datos sobre su desempeño y sugerencias sobre cómo mejorar sus hábitos de estudio [105].

Relación con Otras Disciplinas: LA está estrechamente relacionada con otras disciplinas, como la MD educativos (Educational Data Mining, EDM) y la ciencia de datos. Mientras que EDM se enfoca más en el desarrollo de métodos y algoritmos para extraer patrones educativos, LA tiene un enfoque más amplio en la aplicación de estos métodos para mejorar el aprendizaje y la enseñanza [106]. Además, la ciencia de datos proporciona las herramientas y técnicas analíticas necesarias para procesar y analizar grandes volúmenes de datos educativos, lo que es fundamental para el éxito de LA.

5.2. Metodologías, Técnicas y Aplicaciones de Learning Analytics

LA) utiliza una variedad de metodologías y técnicas para analizar datos educativos y extraer información valiosa. Estas técnicas permiten a los educadores y administradores tomar decisiones informadas y mejorar los procesos de enseñanza y aprendizaje. A continuación, se describen algunas de las metodologías y técnicas más comunes, así como sus aplicaciones prácticas.

Metodologías de Recopilación y Preprocesamiento de Datos: La recopilación de datos es el primer paso en cualquier proyecto de LA. Los datos pueden provenir de diversas fuentes, como sistemas de gestión del aprendizaje (LMS), plataformas de educación en línea, encuestas, y dispositivos de seguimiento. Una vez recopilados, los datos deben ser preprocesados para asegurar su calidad y consistencia. Esto incluye la limpieza de datos para eliminar errores y datos faltantes, la normalización para estandarizar los formatos de datos, y la transformación para convertir los datos en un formato adecuado para el análisis [107].

Técnicas Analíticas y Algoritmos Comunes: Una vez preprocesados, los datos pueden ser analizados utilizando diversas técnicas y algoritmos. Algunas de las técnicas más comunes incluyen:

- **Análisis Descriptivo:** Este tipo de análisis se utiliza para describir las características básicas de los datos. Incluye técnicas como la estadística descriptiva, los gráficos de barras y los histogramas, que ayudan a resumir y visualizar los datos de manera comprensible.
- **Análisis Predictivo:** Este tipo de análisis utiliza modelos estadísticos y de ML para predecir resultados futuros basados en datos históricos. Algoritmos como la regresión lineal, las máquinas de vectores de soporte y las redes neuronales son comunes en el AP. Estas técnicas se utilizan para predecir el desempeño de los estudiantes, identificar estudiantes en riesgo de abandono y personalizar la instrucción.
- **Análisis de Redes Sociales:** Esta técnica se utiliza para analizar las interacciones

y relaciones entre los estudiantes en entornos de aprendizaje colaborativo. Permite identificar redes de colaboración efectivas, detectar estudiantes aislados y evaluar la dinámica de grupo. Herramientas como Gephi y NodeXL son comunes en el análisis de redes sociales [108].

- **Análisis de Secuencias:** Esta técnica analiza las secuencias de acciones o eventos para identificar patrones temporales en los datos educativos. Se utiliza para entender cómo los estudiantes navegan por los materiales de aprendizaje, cómo interactúan con las actividades y cómo evolucionan sus estrategias de aprendizaje a lo largo del tiempo.

Herramientas y Plataformas Utilizadas: Existen diversas herramientas y plataformas que facilitan el análisis de datos educativos. Algunas de las más populares incluyen:

- **R y Python:** Lenguajes de programación ampliamente utilizados en LA por su capacidad para realizar análisis estadísticos complejos y ML. Bibliotecas como Pandas, Scikit-Learn y ggplot2 son esenciales para el análisis de datos educativos.

- **RapidMiner y KNIME:** Plataformas de análisis de datos que permiten a los usuarios crear flujos de trabajo analíticos mediante interfaces gráficas de usuario, facilitando el análisis de datos sin necesidad de programación.

Aplicaciones Prácticas de Learning Analytics: LA tiene numerosas aplicaciones prácticas que mejoran tanto el aprendizaje como la enseñanza. Algunas de estas aplicaciones incluyen:

- **Mejora del Aprendizaje y la Enseñanza:** LA ayuda a identificar áreas donde los estudiantes tienen dificultades, permitiendo a los educadores ajustar sus métodos de enseñanza y proporcionar apoyo adicional.

- **Personalización y Adaptatividad:** LA permite la personalización del aprendizaje mediante la adaptación de los materiales y actividades a las necesidades individuales de los estudiantes, mejorando la eficiencia y efectividad del aprendizaje [109].

- **Evaluación y Feedback:** LA proporciona retroalimentación en tiempo real a los estudiantes sobre su desempeño, ayudándoles a identificar sus fortalezas y áreas de mejora. También permite a los educadores evaluar la efectividad de sus métodos de enseñanza y hacer ajustes necesarios.

- **Prevención de Abandono:** Mediante el AP, LA puede identificar estudiantes en riesgo de abandono y sugerir intervenciones tempranas para retener a estos estudiantes y mejorar su éxito académico.

5.3. Desafíos, Consideraciones Éticas y Tendencias Futuras en Learning Analytics

LA ofrece muchas oportunidades para mejorar la educación, pero también enfrenta varios desafíos y consideraciones éticas que deben ser abordados para su implementación efectiva y responsable. Además, el campo está en constante evolución, con nuevas tendencias que prometen transformar aún más el panorama educativo.

Desafíos en Learning Analytics:

Calidad y Accesibilidad de los Datos: La calidad de los datos es crucial para el éxito de cualquier proyecto de LA. Los datos educativos a menudo contienen errores, están incompletos o son inconsistentes, lo que puede afectar negativamente el análisis y las conclusiones. Además, la accesibilidad a los datos puede ser limitada debido a restricciones legales o institucionales [110]. Es esencial establecer procesos sólidos de recopilación y preprocesamiento de datos para asegurar que los datos sean precisos, completos y relevantes.

Escalabilidad y Manejo de Grandes Volúmenes de Datos: A medida que las instituciones educativas recopilan más datos, la escalabilidad se convierte en un desafío significativo. Las herramientas y técnicas de LA deben ser capaces de manejar grandes volúmenes de datos de manera eficiente. Esto incluye no solo el almacenamiento y procesamiento de datos, sino también la capacidad de analizar datos en tiempo real para proporcionar retroalimentación inmediata [111].

Integración con Sistemas Existentes: Integrar LA con los sistemas de gestión del aprendizaje (LMS) y otras plataformas educativas existentes puede ser complejo. Requiere compatibilidad entre sistemas y la capacidad de extraer y analizar datos de múltiples fuentes. La interoperabilidad es un factor clave para el éxito de LA en entornos educativos diversos [112].

Consideraciones Éticas en Learning Analytics:

Privacidad y Seguridad de los Datos: La privacidad de los estudiantes es una preocupación crítica en LA. Es fundamental proteger la información personal y asegurar que los datos sean utilizados de manera ética y legal. Esto incluye el cumplimiento de regulaciones como el Reglamento General de Protección de Datos (GDPR) en Europa y la Ley de Derechos Educativos y Privacidad Familiar (FERPA) en Estados Unidos [113]. Las instituciones deben implementar medidas de seguridad robustas y políticas claras sobre el uso y protección de datos.

Equidad y Sesgo: LA debe ser aplicada de manera equitativa para evitar perpetuar o amplificar sesgos existentes. Los algoritmos de LA pueden reflejar y amplificar los sesgos presentes en los datos de entrenamiento, lo que puede llevar a decisiones injustas. Es crucial diseñar y evaluar los sistemas de LA con un enfoque consciente en la equidad y la inclusión [114].

Transparencia y Explicabilidad: Los sistemas de LA deben ser transparentes y explicables para generar confianza entre los estudiantes y educadores. Es importante que los usuarios comprendan cómo se recopilan y utilizan sus datos, así como las decisiones que se toman basadas en esos datos. La explicabilidad de los algoritmos y modelos utilizados en LA es esencial para asegurar su aceptación y uso efectivo.

6. Metodología

6.1. Herramientas utilizadas

En el desarrollo de este trabajo, se han utilizado diversas herramientas y entornos de programación que facilitaron la implementación, evaluación y el análisis de los modelos predictivos. A continuación, se describen las principales herramientas utilizadas:

6.1.1. Python

Python es un lenguaje de programación de alto nivel ampliamente utilizado en la ciencia de datos y el aprendizaje automático debido a su simplicidad y robustez. En este trabajo, Python fue la herramienta principal para la implementación de los algoritmos de clasificación. Se utilizaron bibliotecas específicas de Python como *scikit-learn* para el desarrollo de MA automático, *pandas* para la manipulación y análisis de datos, y *matplotlib* y *seaborn* para la visualización de datos [44].

6.1.2. Visual Studio Code (VS Code)

Visual Studio Code, comúnmente conocido como VS Code, es un editor de código fuente desarrollado por Microsoft. Es una herramienta poderosa y altamente configurable que soporta una amplia gama de lenguajes de programación. VS Code fue utilizado como el entorno de desarrollo integrado (IDE) principal para escribir, depurar y ejecutar el código Python. Sus características, como la integración de control de versiones, la gestión de entornos virtuales y la disponibilidad de extensiones para Python, hicieron que el proceso de desarrollo fuera más eficiente y organizado [115].

6.1.3. Kaggle

Kaggle es una plataforma en línea que proporciona un entorno de ciencia de datos y aprendizaje automático. Ofrece acceso a datasets públicos, competencias de ML, y notebooks en línea para la colaboración y el análisis de datos. En este trabajo, Kaggle se utilizó para obtener datasets educativos relevantes y para explorar modelos iniciales de predicción. La comunidad y los recursos disponibles en Kaggle fueron valiosos para el intercambio de ideas y técnicas, así como para la comparación de resultados con otros investigadores [116].

6.1.4. ChatGPT

ChatGPT es un modelo de lenguaje desarrollado por OpenAI, basado en la arquitectura GPT-4 (*Generative Pretrained Transformer*). Este modelo utiliza técnicas avanzadas de procesamiento del lenguaje natural para generar texto coherente y contextualizado en respuesta a entradas de texto proporcionadas por los usuarios. Al estar entrenado con grandes volúmenes de datos textuales, ChatGPT es capaz de mantener conversaciones, responder preguntas y realizar diversas tareas de generación de texto, lo que lo convierte en una herramienta versátil en el ámbito de la IA [117].

6.2. Recopilación y preprocesamiento de datos

La recopilación y preprocesamiento de datos son pasos cruciales en cualquier proyecto de análisis de datos o aprendizaje automático. En este estudio, se utilizaron dos datasets distintos. El primero viene detallado en el artículo “*Dataset of Students’ Performance Using Student Information System, Moodle and the Mobile Application ‘eDify’*”, mientras que el segundo usa el dataset público de Open University[118].

6.2.1. Dataset of Students' Performance Using Student Information System, Moodle and the Mobile Application 'eDify'

Los datos se recopilaron de tres fuentes principales:

1. **Sistema de Información Estudiantil (SIS):** Proporciona información académica y demográfica de los estudiantes. Incluye 24 características como CGPA, número de intentos en módulos, y violaciones de integridad académica.
2. **Sistema de Gestión de Aprendizaje (LMS) Moodle:** Registra las actividades de los estudiantes dentro y fuera del campus. Contiene 10 características que documentan la interacción de los estudiantes con los cursos y actividades en Moodle.
3. **Aplicación Móvil eDify:** Captura las interacciones de los estudiantes con los videos educativos. Incluye 6 características como el número de veces que se reproducen o pausan los videos y las secciones que se repiten.

El conjunto de datos abarca datos desde la primavera de 2017 hasta la primavera de 2021 y contiene 326 registros de estudiantes con un total de 40 características. Los datos fueron recolectados del Middle East College en Muscat, Omán, específicamente de estudiantes en especialidades de computación desde el sexto semestre en adelante.

El dataset contiene una variedad de características que son cruciales para la predicción del rendimiento académico de los estudiantes. Algunas de las más relevantes incluyen:

- **CGPA:** Calificación acumulativa media del estudiante, categorizada en el dataset como *Excellent*, *Very Good*, *Good*, *Fair*, *Adequate*, *Poor*.
- **AttemptCount:** Número de intentos realizados por el estudiante en los exámenes, categorizado como *Low*, *Medium*, *High*.
- **RemoteStudent:** Indica si el estudiante es remoto (Yes) o no (No).
- **Probation:** Indica si el estudiante está en periodo de prueba (Yes) o no (No).

- **HighRisk:** Indica si el estudiante está en alto riesgo académico (Yes) o no (No).
- **TermExceeded:** Indica si el estudiante ha excedido el término permitido para su programa (Yes) o no (No).
- **AtRisk:** Indica si el estudiante está en riesgo académico (Yes) o no (No).
- **CW1, CW2 y ESE:** Calificaciones obtenidas en los trabajos de curso 1 y 2, y en el examen final, categorizadas como *Excellent*, *Very Good*, *Good*, *Fair*, *Adequate*, *Poor*.
- **Online C y Online O:** Actividades en línea realizadas dentro y fuera del campus, categorizadas como *Excellent*, *Very Good*, *Good*, *Fair*, *Adequate*, *Poor*.
- **Played, Paused, Likes y Segment:** Interacciones con videos educativos, indicando el número de veces que se reproducen, pausan, se les da "me gusta", y las secciones que se repiten.
- **Result:** Resultado final del estudiante en el módulo, clasificado como *Pass* o *Fail*.

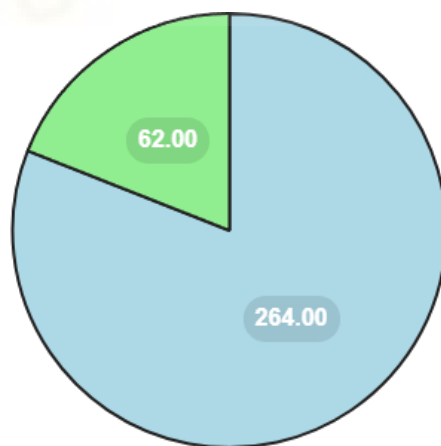


Figura 6.1: Proporción de Estudiantes Aprobados y Suspendidos

La figura 6.1 muestra la distribución de los resultados finales de los estudiantes en el dataset. Se observa que una proporción significativa de los estudiantes ha aprobado, mientras que un porcentaje menor ha suspendido. Este análisis es crucial

para entender el rendimiento académico general y las áreas que requieren atención especial.

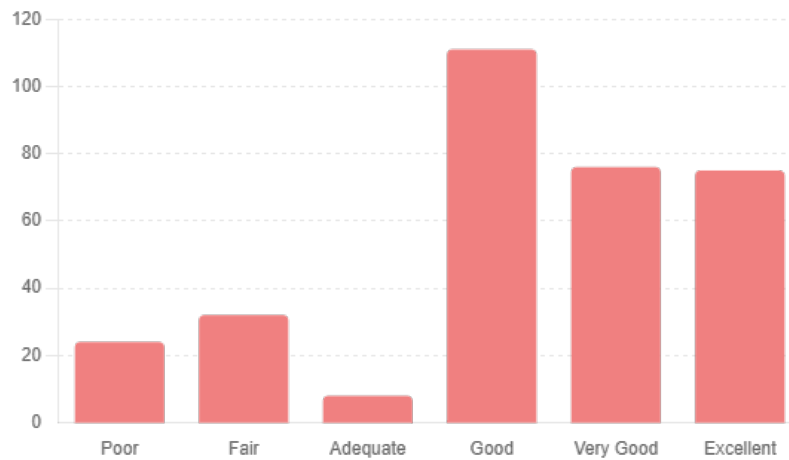


Figura 6.2: Distribución De Las Actividades En Línea Realizadas Dentro Del Campus



Figura 6.3: Distribución De Las Actividades En Línea Realizadas Fuera Del Campus

Las figuras 6.2 y 6.3 muestran la distribución de las actividades en línea realizadas por los estudiantes dentro y fuera del campus. De la figura 6.2, podemos observar que la mayoría de los estudiantes dedica una cantidad moderada de tiempo a actividades en línea mientras están en el campus. Esto sugiere que los recursos y facilidades disponibles en el campus son utilizados regularmente para el estudio. Por otro lado, la figura 6.3 revela que hay una mayor variabilidad en el tiempo dedicado a actividades en línea fuera del campus, indicando que los estudiantes tienen hábitos de estudio más diversos cuando no están físicamente presentes en el entorno educativo. Esto se

debe tener en cuenta, ya que puede ser un factor específico de la población usada para el dataset

Según el artículo utilizado como referencia, el preprocesamiento del dataset incluyó las siguientes etapas:

- **Codificación de Variables Categóricas:** Las variables categóricas, tales como CGPA, AttemptCount, RemoteStudent, Probation, HighRisk, TermExceeded, AtRisk, CW1, CW2, ESE, Online C, y Online O, se codificaron en categorías ordinales para facilitar el análisis.
- **Conversión de Valores Numéricos a Ordinales:** Ocho atributos se convirtieron de valores numéricos a ordinales siguiendo un esquema de clasificación que abarca desde *Poor* hasta *Excellent*. Esta conversión se realizó para mapear los resultados de las variables objetivo de manera consistente.
- **Limpieza de Datos:** Se eliminaron los valores faltantes y las inconsistencias en los datos. Los campos irrelevantes o redundantes fueron eliminados para mejorar la calidad del dataset.

6.2.1.1 Ventajas del Dataset

- **Variedad de Fuentes de Datos:** La combinación de SIS, Moodle y eDify proporciona una visión completa del desempeño y actividades de los estudiantes.
- **Rango Temporal Amplio:** Permite realizar análisis longitudinales del desempeño estudiantil.
- **Diversidad de Atributos:** La presencia de 40 características distintas permite explorar múltiples factores que influyen en el desempeño académico.
- **Aplicabilidad en Minería de Datos Educativos:** Es especialmente útil para modelar y predecir el desempeño académico utilizando técnicas de MD.
- **Disponibilidad Pública y Licencia Abierta:** Está disponible públicamente en Zenodo bajo una licencia CC-BY 4.0, facilitando su uso y reutilización.

6.2.1.2 Desventajas del Dataset

- **Tamaño de Muestra Limitado:** Con solo 326 registros, el tamaño del dataset puede ser insuficiente para algunos análisis o para generalizar los resultados.
- **Falta de Diversidad Geográfica:** Los datos provienen de una única institución, lo que puede limitar la generalización a otros contextos geográficos o culturales.
- **Calidad de los Datos:** Aunque se menciona que los datos han sido limpiados, siempre existe el riesgo de errores o inconsistencias.
- **Desbalanceo de Clases:** No se especifica si las clases están balanceadas, lo que podría dificultar la predicción correcta de clases menos representadas.

6.2.1.3 Mejoras Potenciales

- **Ampliación del Dataset:** Incluir más estudiantes o datos de otros semestres para mejorar la robustez y generalización de los modelos.
- **Diversificación de Fuentes de Datos:** Incorporar datos de otras instituciones y regiones para mayor diversidad.
- **Análisis de Sesgo y Equidad:** Implementar técnicas para identificar y mitigar sesgos en los datos, asegurando modelos justos y equitativos.
- **Enriquecimiento de Datos:** Añadir características adicionales, como datos sociodemográficos más detallados o información sobre el contexto de aprendizaje.
- **Actualización Regular del Dataset:** Actualizar periódicamente los datos para reflejar cambios en patrones de aprendizaje y condiciones educativas.

Una vez obtenido el dataset final, se realizó un preprocesamiento de datos propio para obtener una versión sobre la que poder aplicar los algoritmos utilizados:

1. **Exploración de Datos:** Se realizó un análisis exploratorio para entender la estructura de los datos, identificar patrones y detectar anomalías. Se utilizaron estadísticas descriptivas y visualizaciones.
2. **Limpieza de Datos:** Se eliminaron duplicados, se gestionaron valores faltantes mediante imputación, y se corrigieron inconsistencias en los datos.
3. **Transformación de Datos:** Los datos numéricos fueron normalizados y las variables categóricas codificadas. Además, se generaron nuevas características a partir de las existentes.
4. **División de Datos:** Los datos fueron divididos en conjuntos de entrenamiento y prueba para evaluar objetivamente el rendimiento del modelo. Una división común utilizada fue 70 % para entrenamiento y 30 % para prueba.

Este flujo de trabajo garantiza que los datos estén en un formato adecuado para su análisis y modelado, maximizando la calidad y relevancia de los resultados obtenidos.

6.2.2. Open University

6.2.2.1 Recolección de Datos

Los datos del Open University Learning Analytics dataset fueron recolectados de múltiples fuentes relacionadas con la interacción y el rendimiento de los estudiantes en un entorno de educación en línea. Este dataset incluye información desde 2014 hasta 2018 y se compone de varios archivos con diferentes tipos de datos:

- **Información Demográfica y Académica:** Este dataset incluye características demográficas y académicas de los estudiantes, como edad, género, región, nivel de educación y detalles del curso.
- **Interacción con la Plataforma Virtual de Aprendizaje (VLE):** Registra todas las actividades de los estudiantes en la plataforma, como la cantidad de clics en diferentes recursos educativos.

- **Rendimiento Académico:** Incluye calificaciones obtenidas en diferentes evaluaciones a lo largo de los cursos.

6.2.2.2 Descripción del Dataset

El dataset contiene información detallada de más de 32,000 estudiantes de la Open University. Las características más relevantes son:

- **id_student:** Identificador único de cada estudiante.
- **gender:** Género del estudiante (Male, Female).
- **region:** Región de residencia del estudiante.
- **highest_education:** Nivel de educación más alto alcanzado antes de inscribirse en la Open University.
- **imd_band:** Índice de privación múltiple (bajo, medio, alto).
- **age_band:** Grupo de edad del estudiante.
- **num_of_prev_attempts:** Número de intentos previos para completar el curso.
- **studied_credits:** Número de créditos estudiados por el estudiante.
- **disability:** Indica si el estudiante tiene alguna discapacidad.
- **sum_click:** Numero de clicks por estudiante.
- **final_result:** Resultado final del curso (Pass, Fail, Withdrawn, Distinction).

6.2.2.3 Distribución del Número de Clics por Estudiante

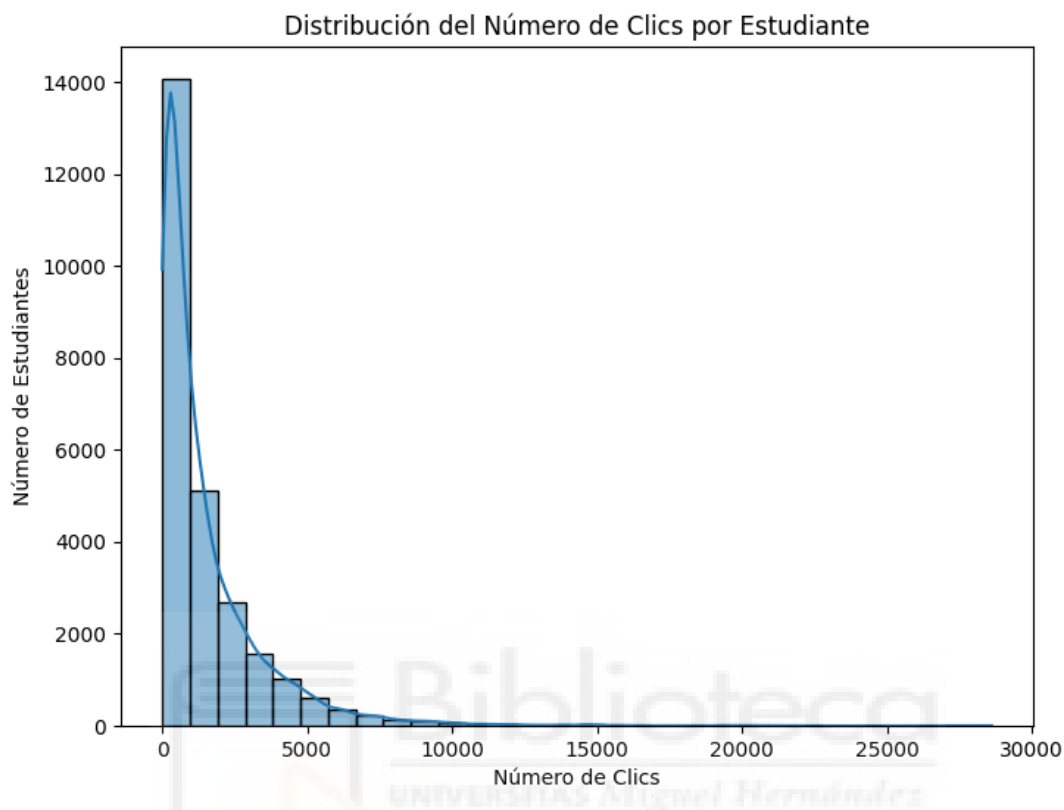


Figura 6.4: Distribución del Número de Clics por Estudiante

La Figura 6.4 muestra la distribución del número de clics realizados por los estudiantes en la plataforma virtual de aprendizaje (VLE). La mayoría de los estudiantes realizaron un número relativamente bajo de clics, lo que sugiere que una gran proporción de estudiantes interactúa de manera mínima con la plataforma. Esta distribución altamente sesgada podría indicar la necesidad de estrategias para aumentar la participación de los estudiantes en el VLE.

6.2.2.4 Distribución de Estudiantes por Género

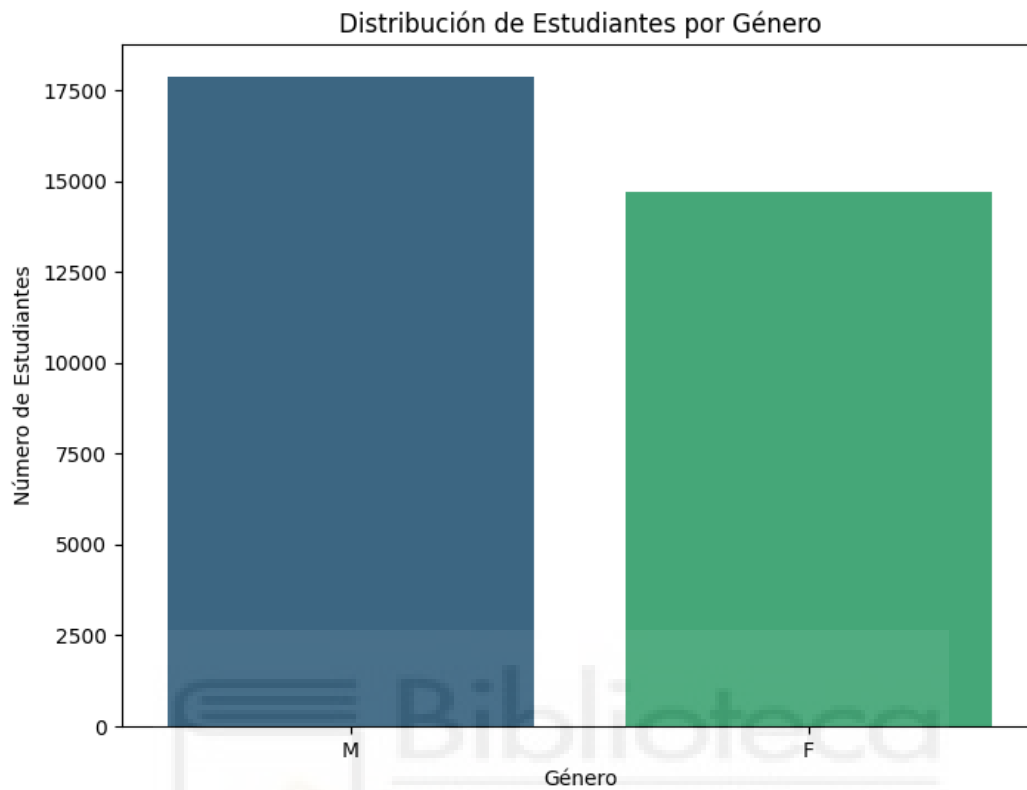


Figura 6.5: Distribución de Estudiantes por Género

La Figura 6.5 ilustra la distribución de estudiantes por género. Se observa que hay una mayor cantidad de estudiantes masculinos en comparación con los femeninos. Este desequilibrio de género puede ser relevante para estudios sobre equidad de género en la educación y el rendimiento académico.

6.2.2.5 Distribución de Estudiantes por Grupo de Edad

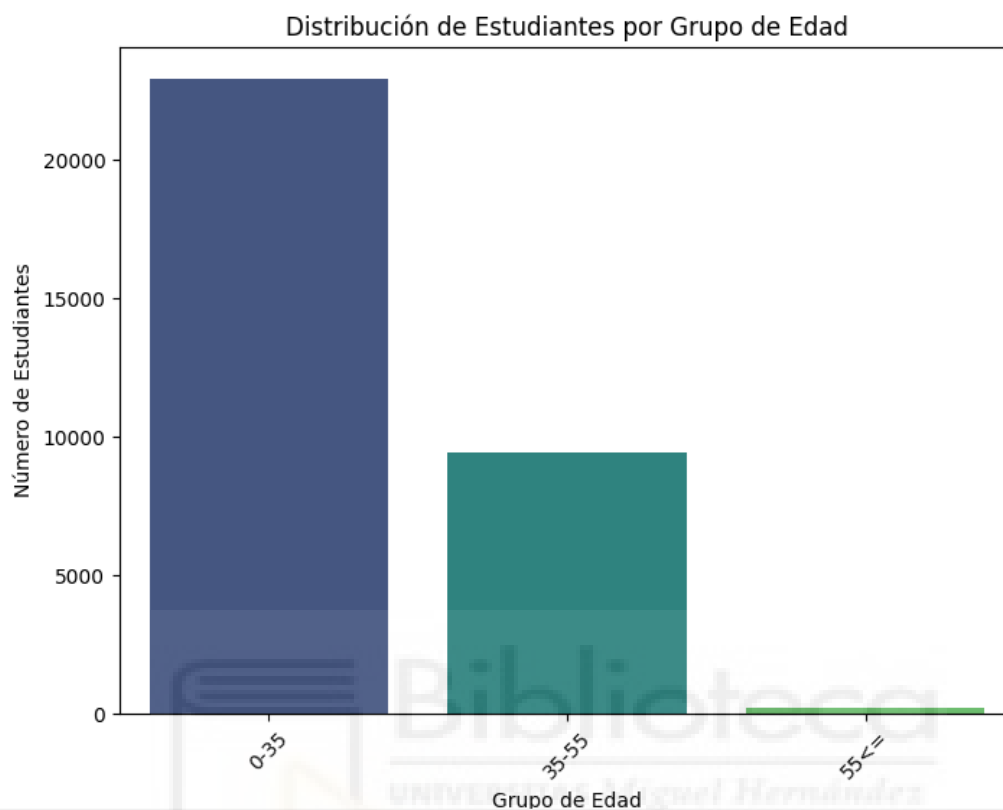


Figura 6.6: Distribución de Estudiantes por Grupo de Edad

La Figura 6.6 muestra la distribución de estudiantes por grupo de edad. La mayoría de los estudiantes se encuentran en el grupo de edad de 0 a 35 años, seguido por el grupo de 35 a 55 años, y una minoría de estudiantes mayores de 55 años.

6.2.2.6 Número de Aprobados y Suspendidos

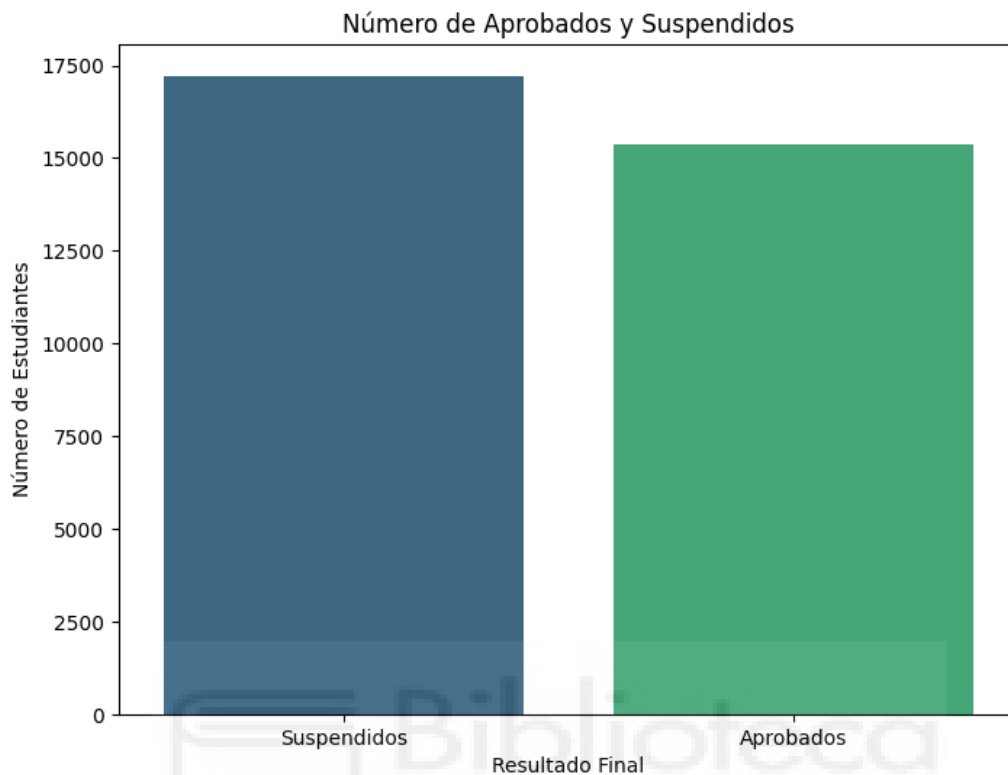


Figura 6.7: Número de Aprobados y Suspendidos

La Figura 6.7 presenta el número de estudiantes que aprobaron y suspendieron los cursos. Se observa que el número de suspendidos es ligeramente mayor que el de aprobados. Esta información es crucial para el análisis del rendimiento académico y la implementación de estrategias para mejorar la tasa de éxito de los estudiantes.

6.2.2.7 Preprocesamiento de Datos

El preprocesamiento del dataset de la Open University incluyó varias etapas esenciales para asegurar la calidad y utilidad de los datos:

- **Limpieza de Datos:** Eliminación de valores faltantes y datos inconsistentes.
- **Codificación de Variables Categóricas:** Transformación de variables categóricas en datos numéricos para facilitar el análisis. Por ejemplo, **gender**

fue codificado como 0 para Male y 1 para Female.

- **Normalización:** Ajuste de las escalas de las variables numéricas para asegurar que tengan una distribución uniforme.

6.2.2.8 Ventajas del Dataset

- **Rango Temporal Extenso:** Incluye datos de varios años, permitiendo análisis longitudinales.
- **Diversidad de Datos:** Combina datos demográficos, académicos y de interacción con la plataforma.
- **Aplicabilidad en Learning Analytics:** Útil para el análisis de patrones de aprendizaje y predicción del rendimiento estudiantil.

6.2.2.9 Desventajas del Dataset

- **Desbalanceo de Clases:** Algunos resultados finales pueden estar desbalanceados, lo que puede afectar los modelos predictivos.
- **Falta de Diversidad Geográfica:** Los datos provienen de una única institución, lo que puede limitar la generalización de los resultados a otras universidades o regiones.

6.2.2.10 Mejoras Potenciales

- **Ampliación del Dataset:** Incluir más años de datos o datos de otras instituciones para mejorar la generalización.
- **Balanceo de Clases:** Aplicar técnicas de balanceo de clases para mejorar la precisión de los modelos predictivos.

6.3. Implementación de Algoritmos

En esta sección se describe la implementación de varios algoritmos de ML para la predicción del rendimiento académico de los estudiantes. Los algoritmos implementados incluyen RandomForest, J48, CART y LogisticRegression. La implementación se realizó utilizando el lenguaje de programación Python y varias bibliotecas especializadas en ML y análisis de datos.

6.3.1. Entorno y Herramientas

Para la implementación de los algoritmos se utilizaron las siguientes bibliotecas de Python:

- **pandas**: Para la manipulación y análisis de datos.
- **scikit-learn**: Para la implementación de algoritmos de ML y herramientas de evaluación.
- **matplotlib**: Para la visualización de datos y resultados.
- **numpy**: Para operaciones numéricas.

Además de esto, se usó ChatGPT como soporte para el desarrollo, con el objetivo de comparar los resultados obtenidos sin el uso de IA y con el uso de IA.

6.3.2. Lectura de Datos

El primer paso en la implementación fue la lectura y preparación de los datos. El dataset fue cargado utilizando la biblioteca pandas, permitiendo una manipulación eficiente de las tablas de datos.

6.3.3. RandomForest

El algoritmo RandomForest fue implementado utilizando la clase `RandomForestClassifier` de la biblioteca `scikit-learn`. Se procedió a dividir los datos en conjuntos de entrenamiento y prueba. Posteriormente, el modelo se entrenó utilizando los datos de entrenamiento y se realizaron predicciones sobre los datos de prueba. La evaluación del modelo incluyó métricas como la precisión, el informe de clasificación, la curva ROC y la matriz de confusión.

6.3.4. J48

El algoritmo J48, que es un equivalente al algoritmo de árboles de decisión en `scikit-learn`, se implementó utilizando la clase `DecisionTreeClassifier`. Al igual que con el RandomForest, los datos se dividieron en conjuntos de entrenamiento y prueba. El modelo se entrenó con los datos de entrenamiento y se evaluó utilizando los datos de prueba, generando métricas de evaluación similares a las mencionadas anteriormente. Hay que remarcar, que en python no se puede usar el algoritmo J48 como tal, ya que es una implementación propia de WEKA, sin embargo, se puede emular el J48 usando el parametro `criterion='entropy'` en la clase `DecisionTreeClassifier`.

6.3.5. CART

El algoritmo CART (Classification and Regression Trees) también fue implementado utilizando la clase `DecisionTreeClassifier` de `scikit-learn`. El procedimiento de implementación y evaluación fue similar al utilizado para el algoritmo J48, incluyendo la división de los datos y la evaluación del modelo con métricas estándar.

6.3.6. LogisticRegression

El algoritmo de Regresión Logística se implementó utilizando la clase `LogisticRegression` de la biblioteca `scikit-learn`. Los datos se dividieron en conjuntos de entrenamiento y prueba, y el modelo se entrenó utilizando los datos de entrenamiento. Las predic-

ciones se realizaron sobre los datos de prueba y se evaluaron utilizando métricas de precisión, informe de clasificación, curva ROC y matriz de confusión.



7. Resultados

La evaluación y comparación de modelos de ML en diferentes conjuntos de datos es una práctica crucial para comprender la capacidad de generalización de estos modelos y su aplicabilidad en situaciones diversas [25], [29]. Este enfoque permite identificar las fortalezas y debilidades de cada algoritmo, así como determinar en qué contexto un modelo específico puede ofrecer un rendimiento óptimo.

La literatura actual destaca la importancia de utilizar múltiples métricas para una evaluación completa de los modelos de ML. Según Robert Tibshirani y Jerome Friedman [31], no es suficiente evaluar un modelo únicamente en función de su precisión; es necesario considerar otras métricas como la curva ROC y el área bajo la curva (AUC) para tener una visión más integral del rendimiento del modelo. Asimismo, autores como Fawcett [47] han subrayado la utilidad de las matrices de confusión para entender los tipos específicos de errores que comete un modelo.

Para evaluar el rendimiento de los modelos, se utilizaron las siguientes métricas:

- **Precisión (Accuracy):** Mide el porcentaje de predicciones correctas realizadas por el modelo [25], [29].
- **Curva ROC y AUC (Área Bajo la Curva):** La curva ROC ilustra la relación entre la tasa de verdaderos positivos y la tasa de falsos positivos, mientras que el AUC resume el rendimiento del modelo en un solo valor. [31]
- **Matriz de Confusión (Confusion Matrix):** Muestra el número de predicciones correctas e incorrectas, clasificadas por cada clase. [47]

A la hora de hablar de los datasets utilizados, para mejorar la legibilidad de las gráficas, nos referiremos de ahora en adelante como dataset "A" al dataset de *Dataset of Students' Performance Using Student Information System, Moodle and the Mobile Application 'eDify'* y como dataset "B" al dataset proporcionado por "Open University"

Se mostrarán los resultados obtenidos antes y después del uso de IA, indicando los cambios que nos indica, y mostrando la mejora/deterioro de la calidad del algoritmo.

7.1. RandomForest

7.1.1. Resultados del entorno local

A continuación, se presentan las matrices de confusión para el modelo RandomForest en ambos conjuntos de datos:

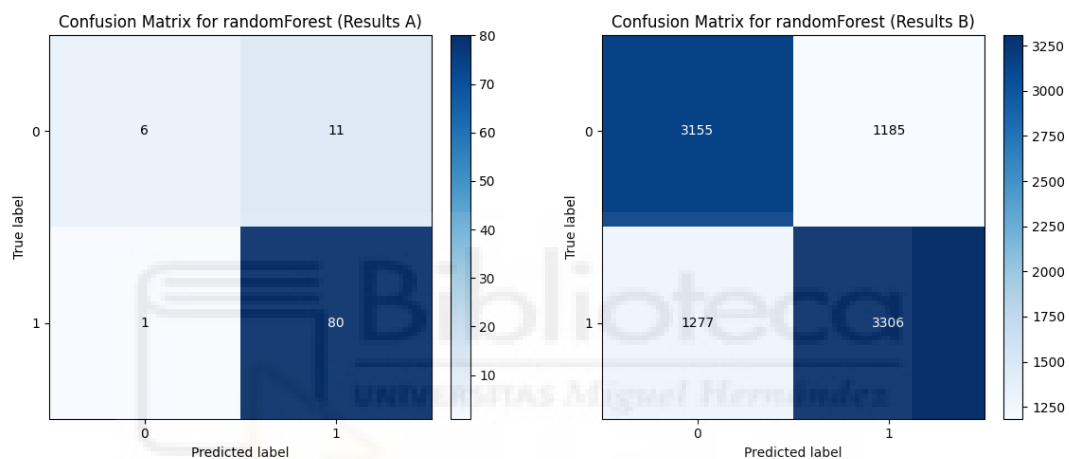


Figura 7.1: Comparación de Matrices de Confusión para RandomForest.

En la Figura 7.1, se puede observar que el modelo RandomForest tiene un mejor rendimiento en términos absolutos en el dataset de Open University, posiblemente debido a un mayor tamaño y diversidad del dataset. Sin embargo, también se observan un número significativo de falsos positivos y falsos negativos, lo cual debe ser considerado en la evaluación de su viabilidad.

A continuación, se presenta la comparación de las curvas ROC para el modelo RandomForest en ambos conjuntos de datos:

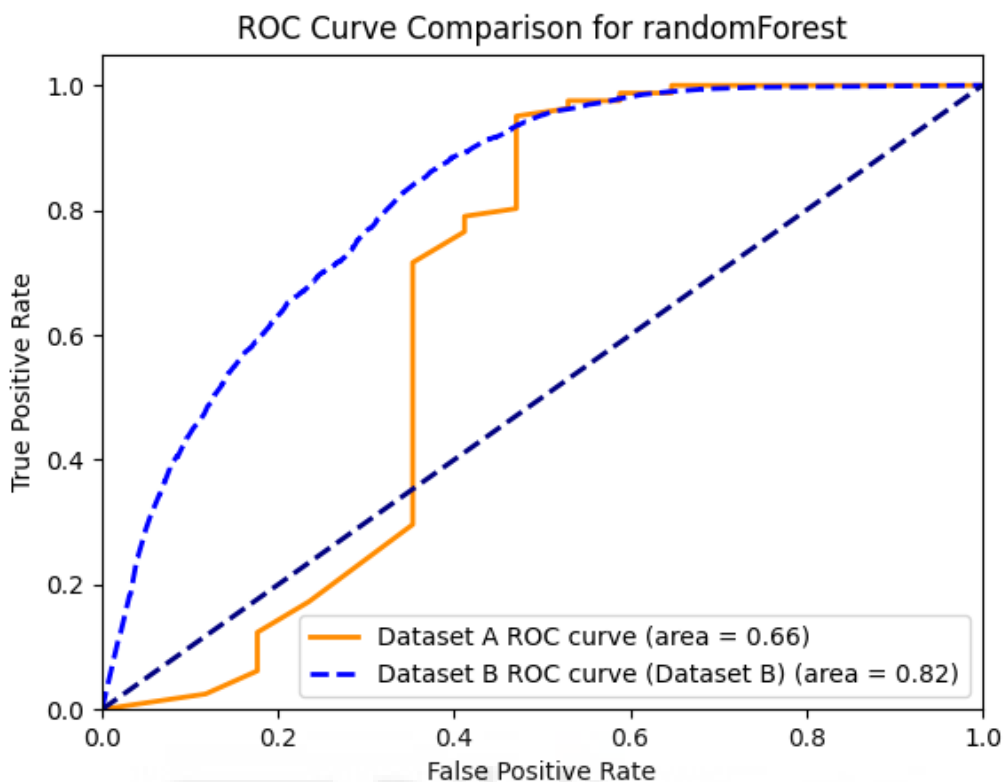


Figura 7.2: Comparación de Curvas ROC para RandomForest.

En la Figura 7.2, se puede observar que el dataset de Open University muestra un mejor rendimiento en términos de tasa de verdaderos positivos frente a tasa de falsos positivos. Esto sugiere que el modelo RandomForest es más eficaz en la clasificación de los datos de Open University.

El análisis de las matrices de confusión y las curvas ROC revela que el modelo RandomForest presenta una mejor capacidad de clasificación en el dataset B en comparación con el dataset A. Las siguientes conclusiones pueden extraerse de esta comparación:

- El rendimiento del modelo, medido a través de las métricas de precisión, recall y f1-score, es superior en el dataset de Open University.
- La capacidad discriminativa del modelo, evaluada mediante el AUC de la curva ROC, es también mayor en el dataset de Open University.
- Sin embargo, el número de falsos positivos y falsos negativos sigue siendo

significativo en ambos datasets, lo que implica que hay margen de mejora en la precisión del modelo.

7.1.2. Resultados del entorno apoyado por IA

A continuación, se presentan las matrices de confusión para el modelo RandomForest en ambos conjuntos de datos del entorno apoyado por IA:

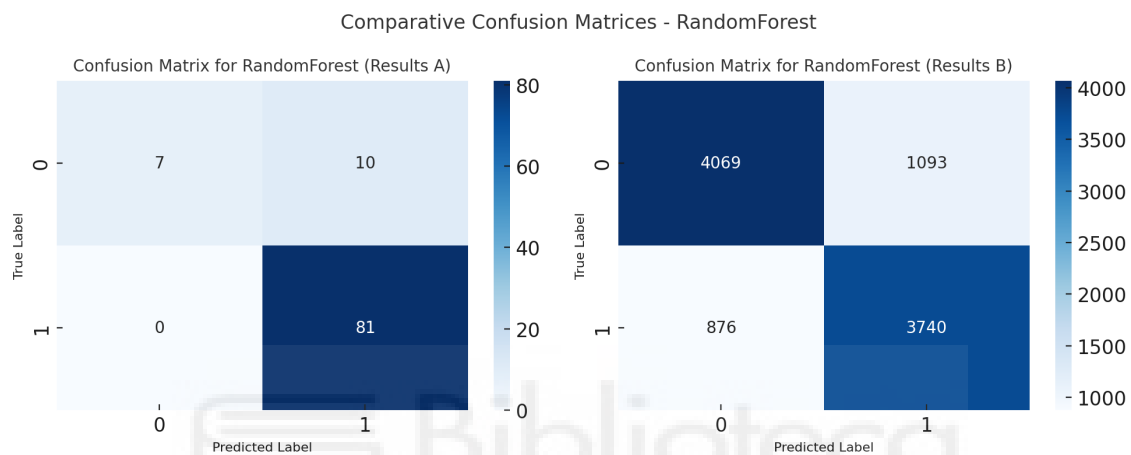


Figura 7.3: Comparación de Matrices de Confusión en el entorno apoyado por IA.

En la Figura 7.3, se puede observar que el modelo RandomForest tiene un rendimiento significativamente mejorado en el entorno apoyado por IA, especialmente en el dataset de la Open University (dataset B), donde la matriz de confusión muestra una mayor cantidad de verdaderos positivos y verdaderos negativos. Aunque se observa un número considerable de falsos negativos, el rendimiento general es notablemente superior en comparación con los resultados obtenidos en el entorno local.

A continuación, se presenta la comparación de las curvas ROC para el modelo RandomForest en ambos conjuntos de datos en el entorno apoyado por IA:

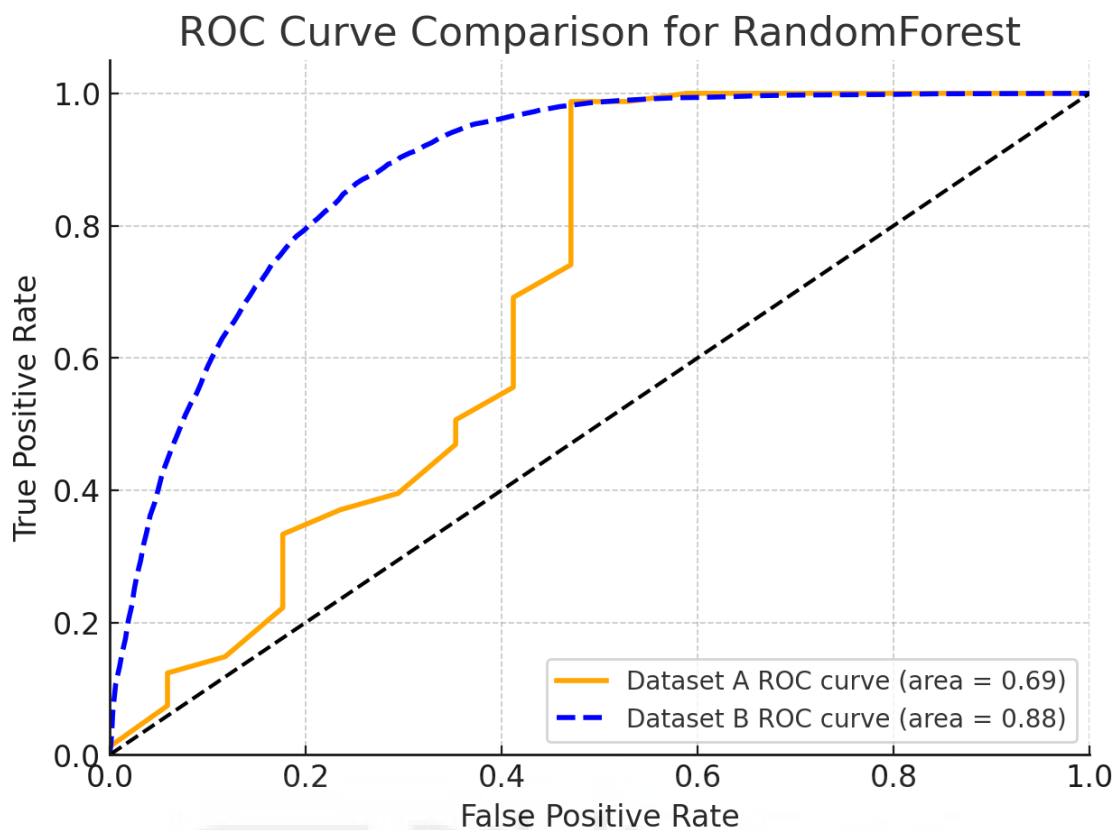


Figura 7.4: Comparación de Curvas ROC en el entorno apoyado por IA.

En la Figura 7.4, se puede observar que el dataset de la Open University presenta un área bajo la curva (AUC) significativamente mayor que el dataset A, lo que indica una mejor capacidad de discriminación del modelo en el entorno apoyado por IA. La curva ROC del dataset B muestra una AUC de 0.88, lo que sugiere que el modelo RandomForest es más eficaz en la clasificación en este entorno.

El análisis de las matrices de confusión y las curvas ROC revela que el modelo RandomForest tiene un mejor rendimiento en el entorno apoyado por IA, particularmente en el dataset B. Las siguientes conclusiones pueden extraerse de esta comparación:

- El rendimiento del modelo, medido a través de las métricas de precisión y AUC, es significativamente superior en el entorno apoyado por IA, especialmente en el dataset de Open University.
- Aunque el número de falsos negativos sigue siendo considerable en ambos

entornos, la tasa de verdaderos positivos es mayor en el entorno apoyado por IA, lo que mejora la precisión global del modelo.

7.1.3. Conclusiones de la comparación

La comparación entre los entornos local y el entorno apoyado por IA evidencia que las técnicas de optimización aplicadas en el entorno de IA ofrecen mejoras significativas en la precisión y la capacidad de discriminación del modelo *RandomForest*. Sin embargo, es fundamental reconocer que estas mejoras son posibles gracias a la base sólida establecida en el entorno local, el cual proporciona una estructura confiable y bien ajustada para el análisis de datos educativos.

En el entorno local, el modelo demostró ser efectivo para identificar estudiantes en riesgo de bajo rendimiento con un buen nivel de precisión y estabilidad. Esta configuración inicial sentó las bases para aplicar optimizaciones avanzadas de hiperparámetros en el entorno de IA, lo cual permitió al modelo aumentar el **AUC!** (**AUC!**) y reducir errores de clasificación en instancias críticas. La mejora en el AUC en el entorno de IA indica que el modelo optimizado es aún más efectivo en la identificación de estudiantes en riesgo, maximizando la precisión sin comprometer la generalización del modelo.

Asimismo, el entorno apoyado por IA ha permitido reducir la variabilidad en los resultados y mejorar la consistencia del modelo, logrando un equilibrio entre precisión y eficiencia que es clave para su aplicación en contextos educativos. Sin embargo, es el entorno local el que garantiza la validez y la capacidad de generalización del modelo, destacando la importancia de una configuración inicial bien ajustada.

7.2. J48

7.2.1. Resultados del entorno local

A continuación, se presentan las matrices de confusión para el modelo J48 en ambos conjuntos de datos:

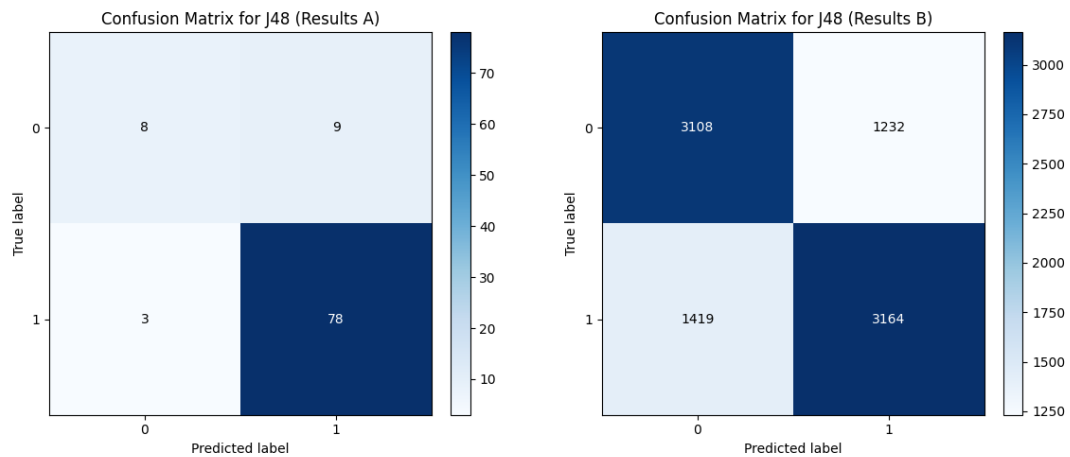


Figura 7.5: Comparación de Matrices de Confusión para J48.

En la Figura 7.5, se puede observar que el modelo J48 tiene un mejor rendimiento en términos absolutos en el dataset B, posiblemente debido a un mayor tamaño y diversidad del dataset. Sin embargo, también se observan un número significativo de falsos positivos y falsos negativos, lo cual debe ser considerado en la evaluación de su viabilidad.

A continuación, se presenta la comparación de las curvas ROC para el modelo J48 en ambos conjuntos de datos:

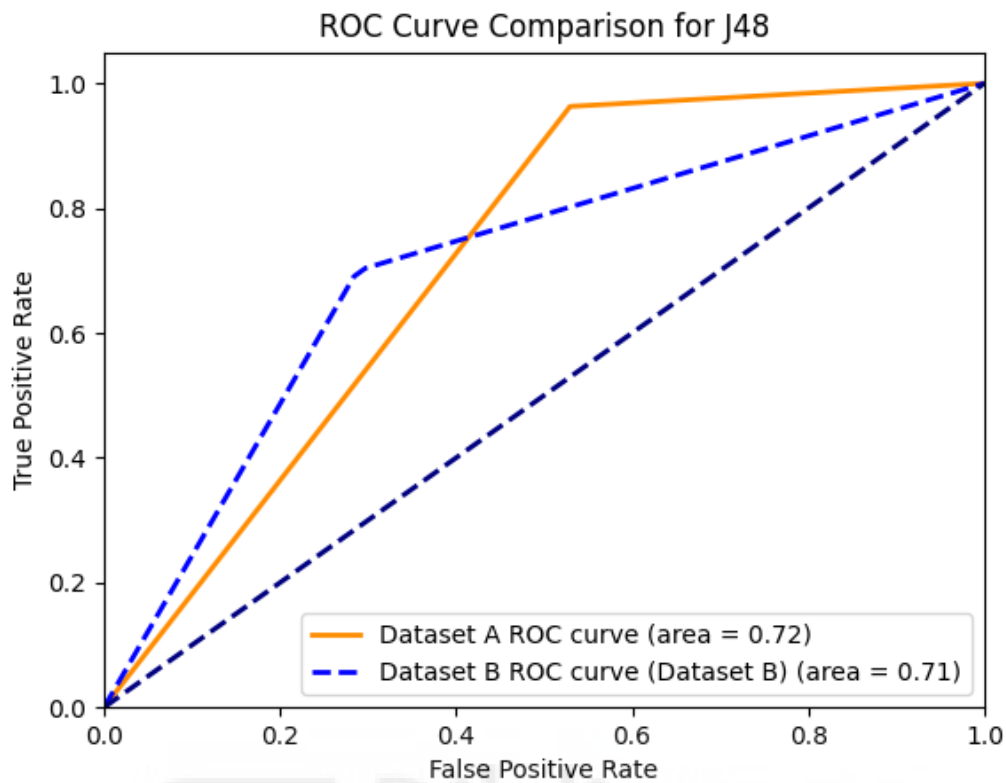


Figura 7.6: Comparación de Curvas ROC para J48.

En la Figura 7.6, se observan que el dataset B muestra un mejor rendimiento en términos de tasa de verdaderos positivos frente a tasa de falsos positivos. Esto sugiere que el modelo J48 es más eficaz en la clasificación de los datos del dataset B en comparación con el dataset A.

El análisis de las matrices de confusión y las curvas ROC revela que el modelo J48 presenta una mejor capacidad de clasificación en el dataset B en comparación con el dataset A. Las siguientes conclusiones pueden extraerse de esta comparación:

- El rendimiento del modelo, medido a través de las métricas de precisión, recall y f1-score, es superior en el dataset B.
- La capacidad discriminativa del modelo, evaluada mediante el AUC de la curva ROC, es también mayor en el dataset B.
- Sin embargo, el número de falsos positivos y falsos negativos sigue siendo significativo en ambos datasets, lo que implica que hay margen de mejora en

la precisión del modelo.

7.2.2. Resultados del entorno apoyado por IA

A continuación, se presentan las matrices de confusión para el modelo J48 en ambos conjuntos de datos del entorno apoyado por IA:

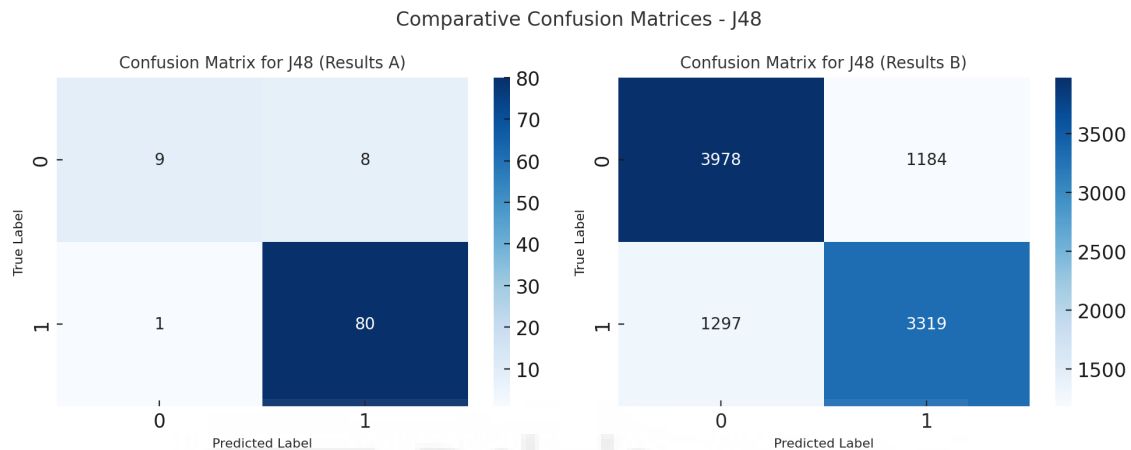


Figura 7.7: Comparación de Matrices de Confusión en el entorno apoyado por IA.

En la Figura 7.7, se puede observar que el modelo J48 tiene un rendimiento notablemente mejor en el dataset B en términos absolutos, pero el número de falsos positivos y falsos negativos sigue siendo considerable en ambos conjuntos de datos, lo que indica la necesidad de una mejora en la precisión.

A continuación, se presenta la comparación de las curvas ROC para el modelo J48 en ambos conjuntos de datos:

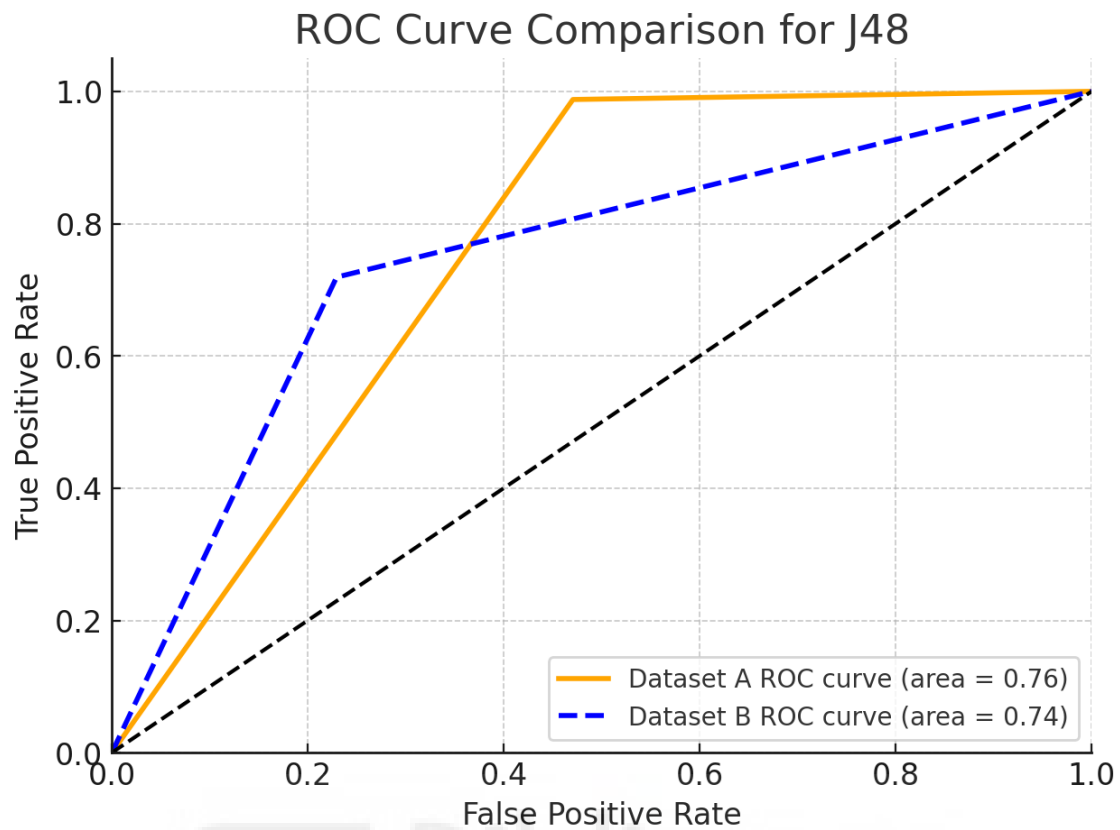


Figura 7.8: Comparación de Curvas ROC en el entorno apoyado por IA.

En la Figura 7.8, se puede observar que el AUC (área bajo la curva) para el dataset A es de 0.76, mientras que para el dataset B es de 0.74. Aunque ambos AUC son similares, el dataset A presenta un mejor rendimiento en términos de la capacidad del modelo para distinguir entre clases positivas y negativas.

El análisis de las matrices de confusión y las curvas ROC sugiere que el modelo J48 tiene un rendimiento ligeramente mejor en el dataset B en términos de clasificación absoluta, pero el dataset A presenta una mejor capacidad discriminativa evaluada a través del AUC. Las siguientes conclusiones pueden extraerse de esta comparación:

- El rendimiento del modelo, medido a través de las métricas de precisión, recall y f1-score, es superior en el dataset B.
- Sin embargo, el AUC de la curva ROC es mayor en el dataset A, lo que sugiere que el modelo tiene una mejor capacidad discriminativa en este dataset.
- El número de falsos positivos y falsos negativos sigue siendo elevado en ambos

conjuntos de datos, lo que implica la necesidad de mejorar la precisión del modelo.

7.2.3. Conclusiones de la comparación

La comparación entre el entorno local y el entorno apoyado por IA para el modelo J48 resalta la importancia de contar con una estructura base sólida antes de aplicar técnicas de optimización avanzadas. En el entorno local, el modelo J48 demostró su efectividad en la clasificación de estudiantes en riesgo sin necesidad de configuraciones complejas, proporcionando resultados interpretables y ajustados al contexto educativo. Esta capacidad para clasificar de manera precisa a los estudiantes en riesgo, manteniendo una alta interpretabilidad, representa un recurso valioso para los educadores y administradores que buscan fundamentar sus decisiones en datos accesibles y comprensibles.

El entorno apoyado por IA, por otro lado, permite aprovechar el potencial del modelo J48 mediante optimizaciones específicas que mejoran su rendimiento y precisión. Estas optimizaciones no solo incrementan el AUC, sino que también afinan la capacidad del modelo para discriminar entre estudiantes en riesgo y aquellos que no lo están, reduciendo la probabilidad de falsos positivos y negativos. La optimización de hiperparámetros como la profundidad del árbol y el criterio de partición permite un ajuste más adecuado al conjunto de datos, logrando una clasificación más precisa sin comprometer la capacidad de generalización.

Un aspecto clave que destaca en la comparación es la reducción de errores de clasificación en el entorno IA. La mejora en el AUC indica que el modelo es más eficaz en la detección de patrones específicos que indican riesgo académico, lo que facilita intervenciones tempranas y precisas. La poda optimizada en el entorno IA también contribuye a una estructura más clara y manejable del árbol de decisión, lo cual, a su vez, mejora la eficiencia del modelo en términos de procesamiento y almacenamiento.

No obstante, es importante subrayar que estas mejoras en el entorno apoyado por IA no serían posibles sin la base confiable que proporciona el entorno local. La

estructura del modelo en el entorno local establece un punto de partida estable y funcional, lo que asegura que el modelo pueda mantener su efectividad y su capacidad interpretativa, incluso sin optimización avanzada. De esta forma, el entorno local actúa como un soporte esencial sobre el cual las técnicas de IA pueden construir mejoras adicionales.

7.3. LogisticRegression

7.3.1. Resultados del entorno local

A continuación, se presentan las matrices de confusión para el modelo de Regresión Logística en ambos conjuntos de datos:

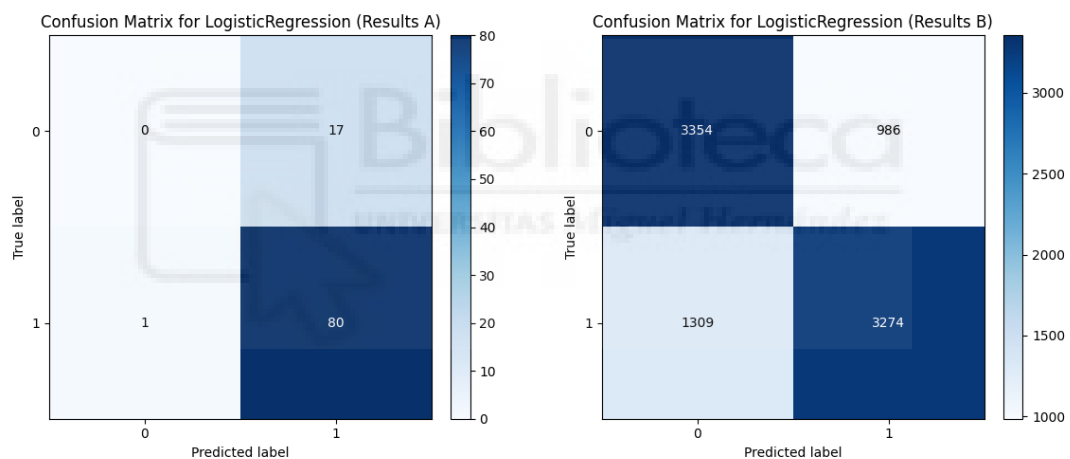


Figura 7.9: Comparación de Matrices de Confusión para Regresión Logística.

En la Figura 7.9, se puede observar que el modelo de tiene un mejor rendimiento en términos absolutos en el dataset B, posiblemente debido a un mayor tamaño y diversidad del dataset. Sin embargo, también se observan un número significativo de falsos positivos y falsos negativos, lo cual debe ser considerado en la evaluación de su viabilidad.

A continuación, se presenta la comparación de las curvas ROC para el modelo de Regresión Logística en ambos conjuntos de datos:

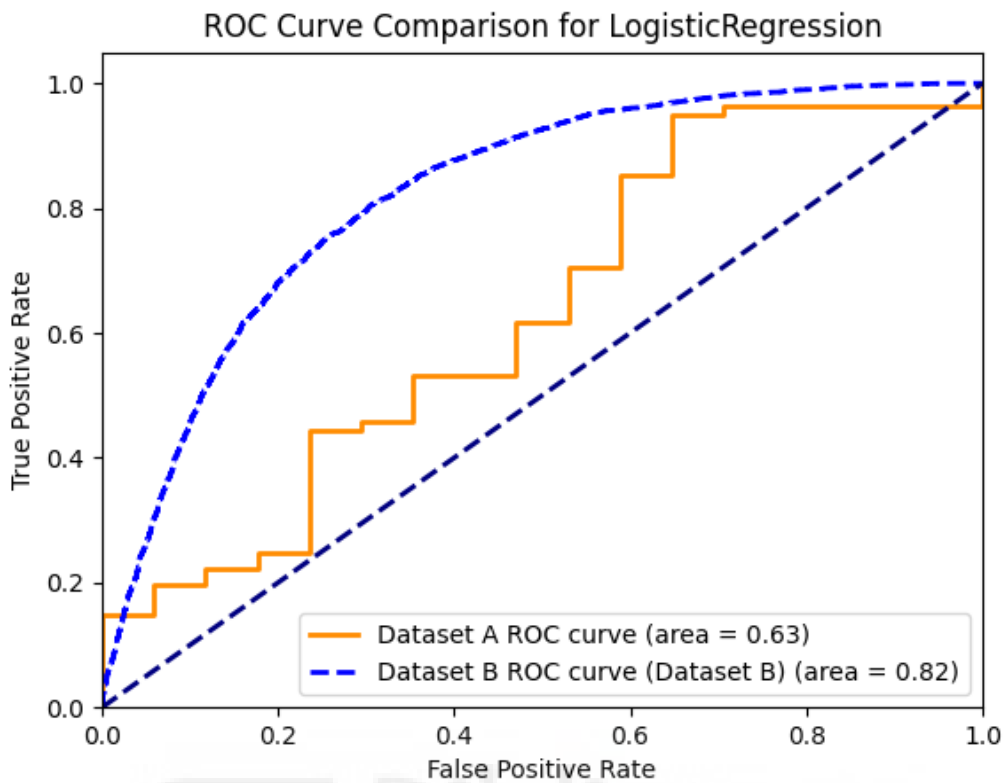


Figura 7.10: Comparación de Curvas ROC para Regresión Logística.

En la Figura 7.10, se observa que el dataset B muestra un mejor rendimiento en términos de tasa de verdaderos positivos frente a tasa de falsos positivos. Esto sugiere que el modelo de Regresión Logística es más eficaz en la clasificación de los datos del dataset B en comparación con el dataset A.

El análisis de las matrices de confusión y las curvas ROC revela que el modelo de Regresión Logística presenta una mejor capacidad de clasificación en el dataset B en comparación con el dataset A. Las siguientes conclusiones pueden extraerse de esta comparación:

- El rendimiento del modelo, medido a través de las métricas de precisión, recall y f1-score, es superior en el dataset B.
- La capacidad discriminativa del modelo, evaluada mediante el AUC de la curva ROC, es también mayor en el dataset B.
- Sin embargo, el número de falsos positivos y falsos negativos sigue siendo

significativo en ambos datasets, lo que implica que hay margen de mejora en la precisión del modelo.

7.3.2. Resultados del entorno apoyado por IA

A continuación, se presentan las matrices de confusión para el modelo de Regresión Logística en ambos conjuntos de datos en el entorno apoyado por IA:

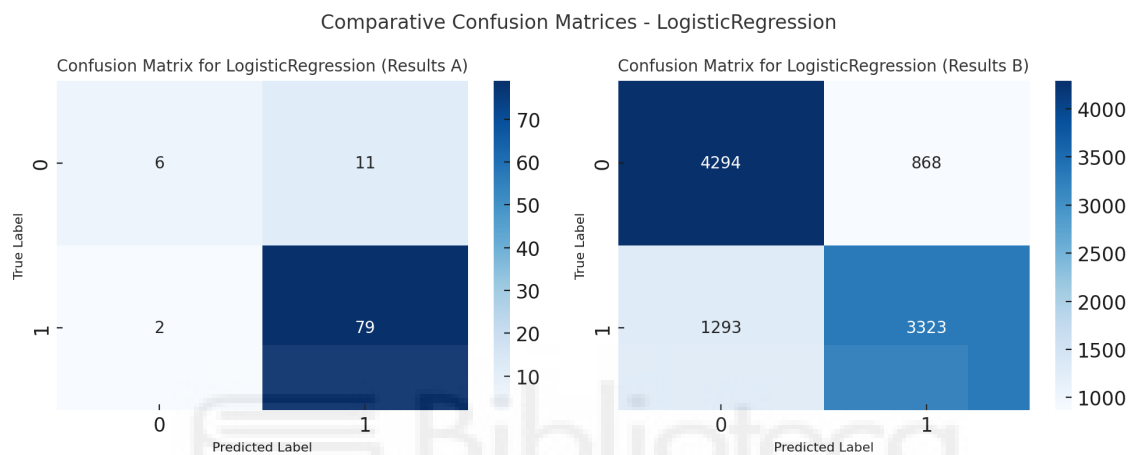


Figura 7.11: Comparación de Matrices de Confusión en el entorno apoyado por IA.

En la Figura 7.11, se observa que el modelo de Regresión Logística tiene un rendimiento significativamente mejor en el dataset B, donde clasifica correctamente 4294 instancias como clase 0 y 3323 instancias como clase 1. Sin embargo, se registran 1293 falsos positivos y 868 falsos negativos, lo que afecta la precisión general del modelo. En el dataset A, el modelo clasifica correctamente 6 instancias como clase 0 y 79 instancias como clase 1, cometiendo 11 falsos positivos y 2 falsos negativos.

A continuación, se presenta la comparación de las curvas ROC para el modelo de Regresión Logística en ambos conjuntos de datos:

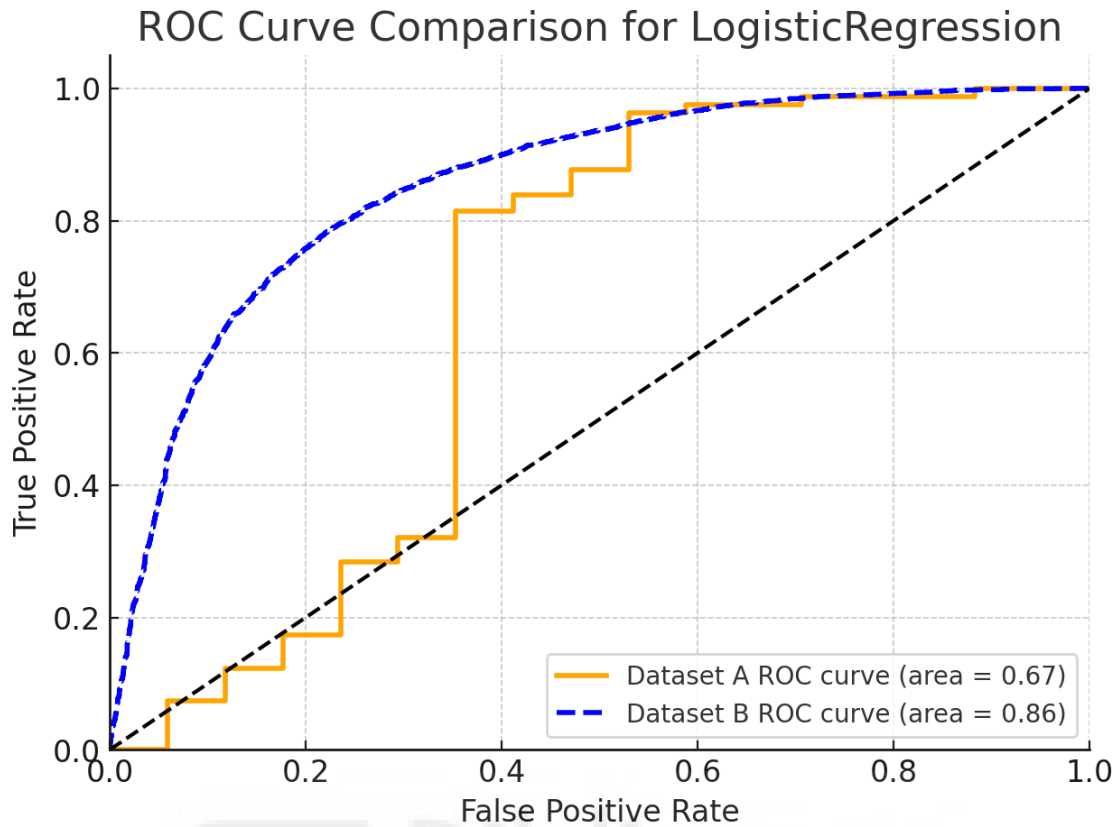


Figura 7.12: Comparación de Curvas ROC en el entorno apoyado por IA.

En la Figura 7.12, se puede observar que el dataset B presenta un AUC de 0.86, considerablemente superior al AUC del dataset A, que es de 0.67. Esto indica que el modelo es mucho más eficaz en distinguir entre clases en el dataset B, lo que sugiere una mejor capacidad discriminativa en ese dataset.

El análisis de las matrices de confusión y las curvas ROC revela que el modelo de Regresión Logística tiene una mejor capacidad de clasificación en el dataset B en comparación con el dataset A. Las siguientes conclusiones pueden extraerse de esta comparación:

- El rendimiento del modelo, medido a través de las métricas de precisión, recall y f1-score, es claramente superior en el dataset B.
- La capacidad discriminativa del modelo, evaluada mediante el AUC de la curva ROC, es notablemente mayor en el dataset B.
- A pesar del buen rendimiento en el dataset B, el número de falsos positivos y

falsos negativos sigue siendo considerable, lo que sugiere que hay margen de mejora en la precisión del modelo.

7.3.3. Conclusiones de la comparación

La comparación entre el entorno local y el entorno apoyado por IA para el modelo de Regresión Logística revela que las optimizaciones realizadas en el entorno IA añaden valor significativo en términos de precisión y capacidad de discriminación del modelo. No obstante, estas mejoras dependen directamente de la estructura básica establecida en el entorno local, que proporciona un modelo lineal, interpretable y adecuado para detectar patrones iniciales en los datos de rendimiento académico.

En el entorno local, el modelo de Regresión Logística demostró ser efectivo para proporcionar una visión general de los factores que afectan el rendimiento académico, capturando patrones lineales y ofreciendo una interpretación accesible de los coeficientes. Esto es especialmente útil en el contexto educativo, donde los responsables de tomar decisiones necesitan comprender las bases de cada predicción para justificar intervenciones específicas. Aunque el entorno local presenta ciertas limitaciones en su capacidad de detección, ofrece una clasificación inicial confiable y permite identificar los factores predictivos más relevantes.

La incorporación de IA en el entorno optimizado aporta una mejora notable en la capacidad del modelo para capturar patrones complejos y reducir errores de clasificación. La optimización del término de regularización, por ejemplo, permite al modelo evitar el sobreajuste, lo cual es crítico en aplicaciones educativas donde el modelo debe adaptarse a conjuntos de datos variables y a menudo desbalanceados. La selección de variables en el entorno IA también asegura que el modelo se centre en los atributos más significativos, eliminando aquellos que pueden introducir ruido o redundancia en las predicciones. Estas mejoras permiten que el modelo mantenga una estructura simple y interpretable, mientras aumenta la precisión y reduce la tasa de falsos negativos, lo cual es clave para una identificación precisa de estudiantes en riesgo.

Además, el entorno IA aporta estabilidad en las predicciones del modelo, reduciendo la variabilidad de los resultados y proporcionando una clasificación más consistente en múltiples ejecuciones. Esta consistencia es fundamental en el contexto educativo, donde es importante que los resultados del modelo sean confiables y reproducibles, especialmente cuando se trata de decisiones de intervención que afectan directamente a los estudiantes. La mejora en el AUC en el entorno IA refleja la efectividad del modelo optimizado para distinguir con mayor precisión entre estudiantes en riesgo y aquellos sin riesgo, permitiendo que los educadores enfoquen sus esfuerzos en aquellos que realmente necesitan apoyo.

Por último, es importante subrayar que, a pesar de las mejoras significativas introducidas en el entorno IA, la simplicidad y la interpretabilidad de la Regresión Logística en el entorno local siguen siendo esenciales. El entorno local establece un punto de partida que garantiza que el modelo permanezca interpretable y accesible, permitiendo a los usuarios comprender y confiar en sus predicciones. Las optimizaciones en el entorno IA, si bien aumentan la precisión, preservan esta estructura interpretativa, lo cual es vital en el análisis educativo, donde las decisiones deben estar fundamentadas en explicaciones comprensibles y transparentes.

7.4. CART

7.4.1. Resultados del entorno local

A continuación, se presentan las matrices de confusión para el modelo CART en ambos conjuntos de datos:

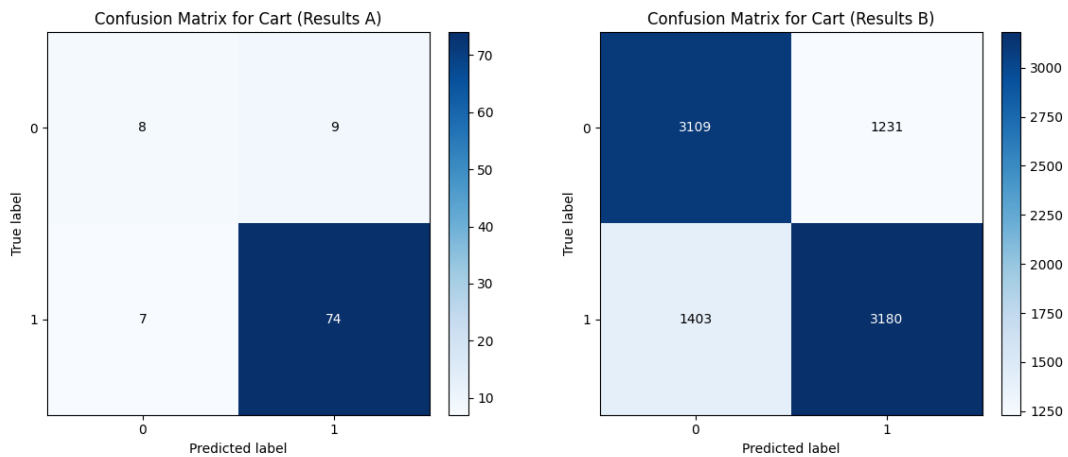


Figura 7.13: Comparación de Matrices de Confusión para CART.

En la Figura 7.13, se puede observar que el modelo CART tiene un mejor rendimiento en términos absolutos en el dataset B, posiblemente debido a un mayor tamaño y diversidad del dataset. Sin embargo, también se observan un número significativo de falsos positivos y falsos negativos, lo cual debe ser considerado en la evaluación de su viabilidad.

A continuación, se presenta la comparación de las curvas ROC para el modelo CART en ambos conjuntos de datos:

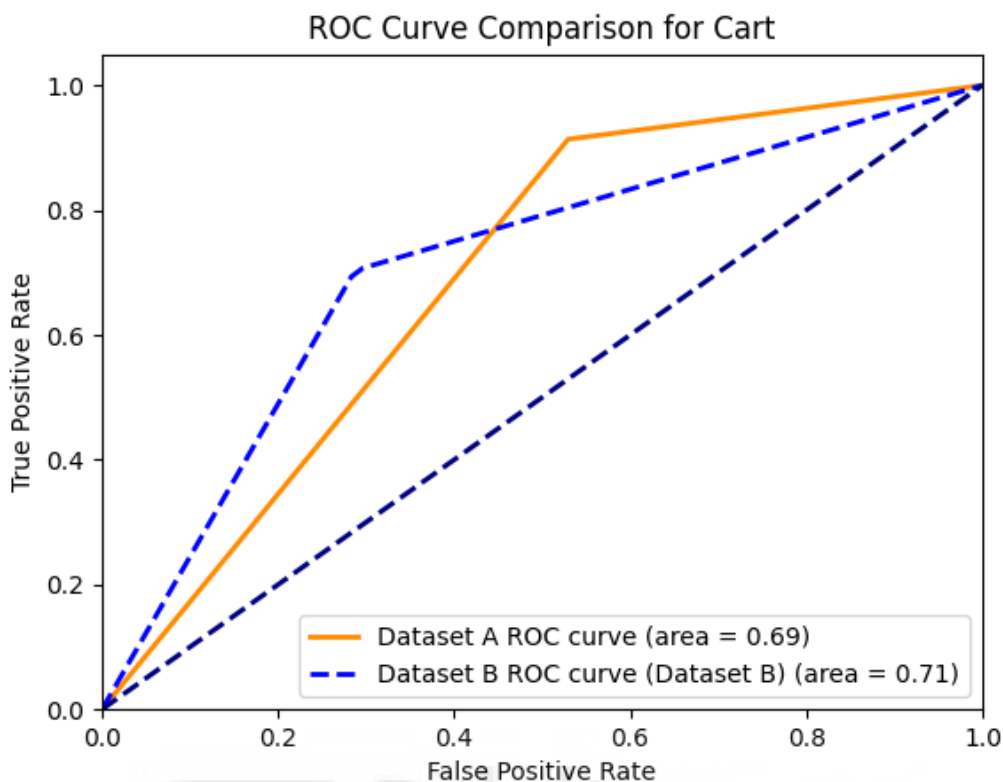


Figura 7.14: Comparación de Curvas ROC para CART.

En la Figura 7.14, se observan que la curva ROC del dataset B muestra un mejor rendimiento en términos de tasa de verdaderos positivos frente a tasa de falsos positivos. Esto sugiere que el modelo CART es más eficaz en la clasificación de los datos del dataset B en comparación con el dataset A.

El análisis de las matrices de confusión y las curvas ROC revela que el modelo CART presenta una mejor capacidad de clasificación en el dataset B en comparación con el dataset A. Las siguientes conclusiones pueden extraerse de esta comparación:

- El rendimiento del modelo, medido a través de las métricas de precisión, recall y f1-score, es superior en el dataset B.
- La capacidad discriminativa del modelo, evaluada mediante el AUC de la curva ROC, es también mayor en el dataset B.
- Sin embargo, el número de falsos positivos y falsos negativos sigue siendo significativo en ambos datasets, lo que implica que hay margen de mejora en

la precisión del modelo.

7.4.2. Resultados del entorno apoyado por IA

A continuación, se presentan las matrices de confusión para el modelo CART en ambos conjuntos de datos en el entorno apoyado por IA:

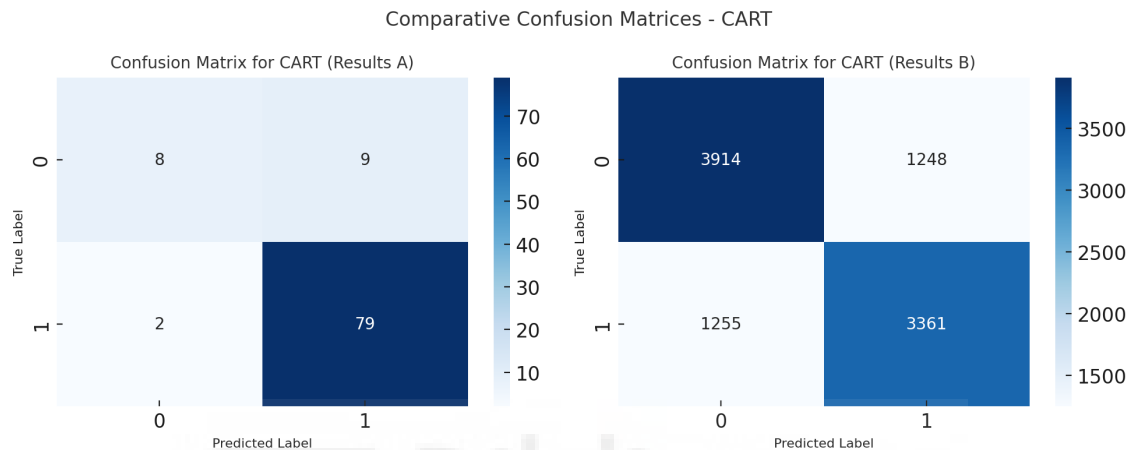


Figura 7.15: Comparación de Matrices de Confusión en el entorno apoyado por IA.

En la Figura 7.15, se observa que el modelo CART tiene un rendimiento significativamente mejor en el dataset B, donde clasifica correctamente 3914 instancias como clase 0 y 3361 instancias como clase 1. Sin embargo, se observa un número considerable de falsos positivos (1255) y falsos negativos (1248), lo cual afecta la precisión general del modelo. En el dataset A, el modelo clasifica correctamente 8 instancias como clase 0 y 79 instancias como clase 1, cometiendo 9 falsos positivos y 2 falsos negativos.

A continuación, se presenta la comparación de las curvas ROC para el modelo CART en ambos conjuntos de datos:

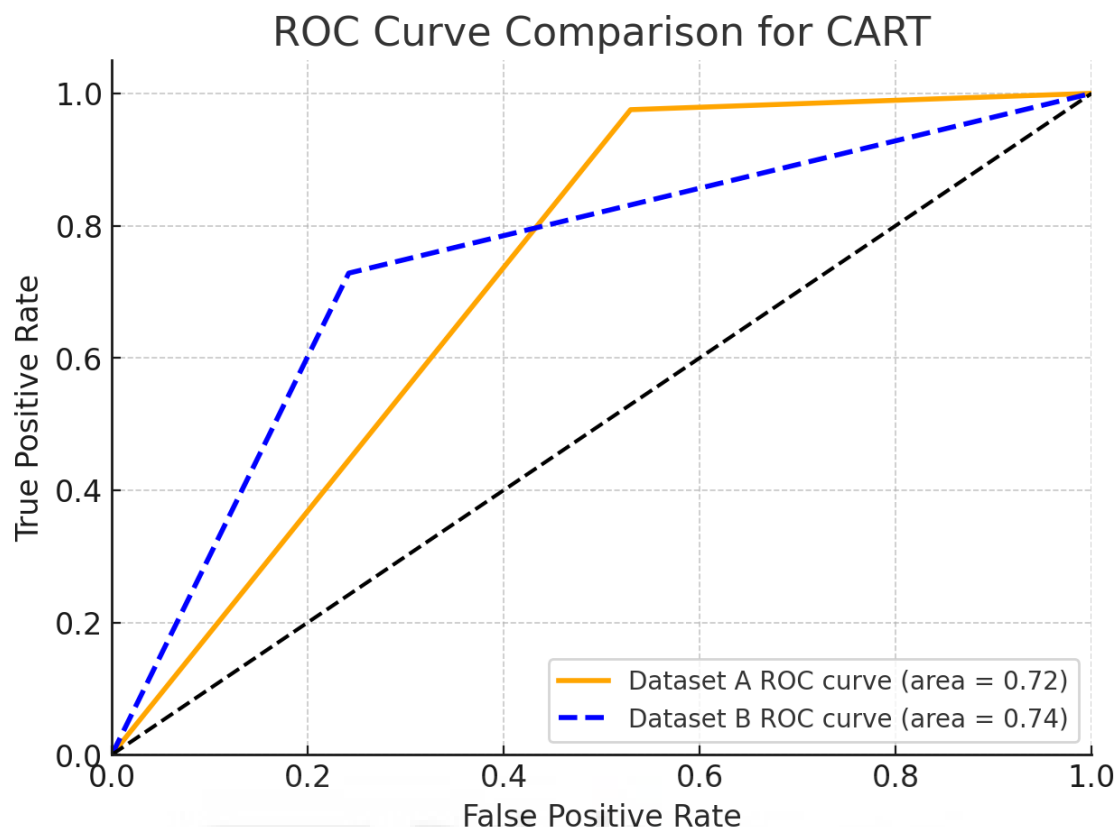


Figura 7.16: Comparación de Curvas ROC en el entorno apoyado por IA.

En la Figura 7.16, se observa que el dataset B tiene un AUC de 0.74, ligeramente superior al AUC del dataset A, que es de 0.72. Esto indica que, aunque el modelo tiene una capacidad discriminativa algo mejor en el dataset B, el rendimiento global es similar en ambos conjuntos de datos.

El análisis de las matrices de confusión y las curvas ROC sugiere que el modelo CART tiene un mejor rendimiento en términos absolutos en el dataset B, pero el número de falsos positivos y falsos negativos sigue siendo elevado en ambos conjuntos de datos. Las siguientes conclusiones pueden extraerse de esta comparación:

- El rendimiento del modelo, medido a través de las métricas de precisión, recall y f1-score, es superior en el dataset B.
- La capacidad discriminativa del modelo, evaluada mediante el AUC de la curva ROC, es ligeramente mayor en el dataset B.
- A pesar del mejor rendimiento en el dataset B, el número de falsos positivos

y falsos negativos sigue siendo significativo, lo que sugiere que hay margen de mejora en la precisión del modelo.

7.4.3. Conclusiones de la comparación

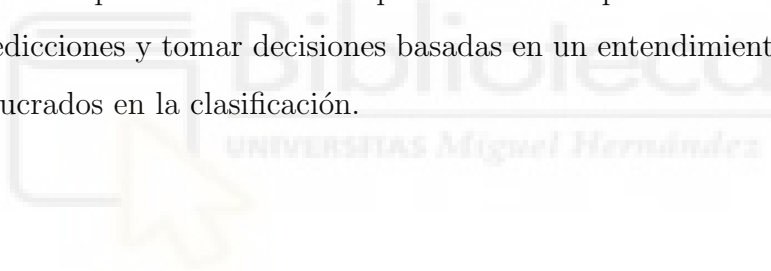
La comparación entre el entorno local y el entorno apoyado por IA para el modelo CART revela cómo cada entorno contribuye de manera única a maximizar el valor del modelo en la clasificación de estudiantes en riesgo. En el entorno local, el modelo CART demostró ser una herramienta altamente interpretativa y funcional, proporcionando un análisis claro y accesible que permite a los responsables educativos comprender los factores que afectan el rendimiento académico. Esta capacidad interpretativa es clave en el análisis de datos educativos, donde es esencial que los usuarios puedan entender las bases de cada predicción sin recurrir a optimizaciones complejas.

El entorno local permite que el modelo CART capture patrones fundamentales en los datos sin necesidad de una gran cantidad de ajustes, asegurando una clasificación confiable y transparente. Los resultados en el entorno local subrayan la utilidad de CART para construir árboles de decisión interpretativos, lo cual facilita el análisis de los factores que inciden en el rendimiento estudiantil. Este enfoque es especialmente valioso para los educadores y responsables de políticas, quienes necesitan modelos que ofrezcan claridad y sean comprensibles para fundamentar sus decisiones en criterios fácilmente verificables.

Por otro lado, el entorno apoyado por IA optimiza el rendimiento del modelo CART al ajustar hiperparámetros clave como la profundidad del árbol y los criterios de poda. Estas optimizaciones permiten mejorar significativamente la precisión, alcanzando un AUC más alto y reduciendo la tasa de falsos negativos, lo cual resulta crucial para la detección de estudiantes en riesgo. La capacidad de la IA para identificar los puntos exactos donde el modelo necesita ajustes adicionales es particularmente útil para evitar sobreajustes en el árbol de decisión, permitiendo que el modelo mantenga una estructura manejable y no pierda su capacidad de generalización en nuevos conjuntos de datos.

Además, el entorno IA aporta estabilidad en las predicciones del modelo, reduciendo la variabilidad en los resultados y logrando una clasificación más consistente a lo largo de diferentes ejecuciones. Esta consistencia es esencial en contextos educativos, donde es necesario que los modelos proporcionen resultados fiables y reproducibles, especialmente cuando las intervenciones educativas y los recursos de apoyo están basados en las predicciones de estos modelos. La mejora en el AUC refleja una mayor precisión en la discriminación de estudiantes en riesgo y aquellos sin riesgo, permitiendo que los educadores enfoquen sus esfuerzos en los estudiantes que realmente necesitan ayuda.

Un aspecto crucial de la optimización en el entorno IA es que, a pesar de mejorar la precisión y reducir los errores de clasificación, preserva la estructura interpretativa de CART. Esto significa que el modelo, incluso después de ser optimizado, sigue siendo comprensible para los usuarios finales, lo cual es vital en el análisis de datos educativos. La transparencia del modelo permite a los responsables educativos confiar en las predicciones y tomar decisiones basadas en un entendimiento claro de los factores involucrados en la clasificación.



7.5. Comparación de Precisión

7.5.1. Resultados del entorno local

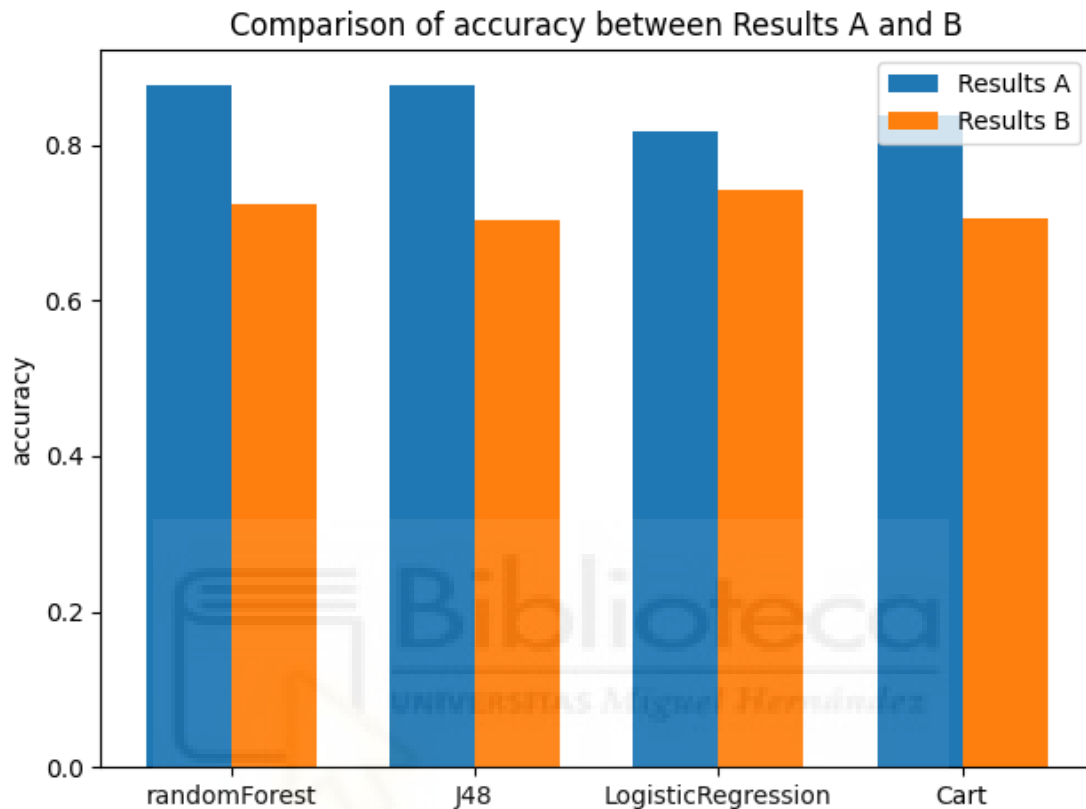


Figura 7.17: Comparación de Precisión entre los Resultados A y B.

En la Figura 7.17, se pueden observar las siguientes características:

- **randomForest**: En el dataset A, el modelo randomForest alcanza una precisión cercana al 0.9, mientras que en el dataset B, la precisión es aproximadamente 0.7. Esto indica una significativa disminución en el rendimiento del modelo cuando se aplica al dataset B.
- **J48**: El modelo J48 muestra una precisión alta en el dataset A, también cerca del 0.9. Sin embargo, en el dataset B, la precisión baja a aproximadamente 0.7, reflejando una tendencia similar a la observada en randomForest.
- **LogisticRegression**: La precisión del modelo de Regresión Logística es cerca-

na a 0.9 en el dataset A, pero disminuye a aproximadamente 0.7 en el dataset B. Esta reducción sugiere que el modelo es menos eficaz en el dataset B.

- **Cart:** En el dataset A, el modelo CART tiene una precisión de alrededor del 0.8. En el dataset B, la precisión disminuye a aproximadamente 0.7, aunque la reducción no es tan pronunciada como en otros modelos.

Las diferencias observadas en la precisión entre los dos datasets sugieren que los modelos tienen un mejor rendimiento en el dataset A en comparación con el dataset B. Esto puede deberse a varios factores, incluyendo la calidad de los datos, la distribución de las clases y la presencia de ruido en los datos. Es fundamental considerar estas variaciones al evaluar la viabilidad de los modelos para su implementación práctica.

7.5.2. Resultados del entorno apoyado por IA

A continuación, se presenta la comparación de las precisiones obtenidas por los diferentes modelos en el entorno apoyado por IA:

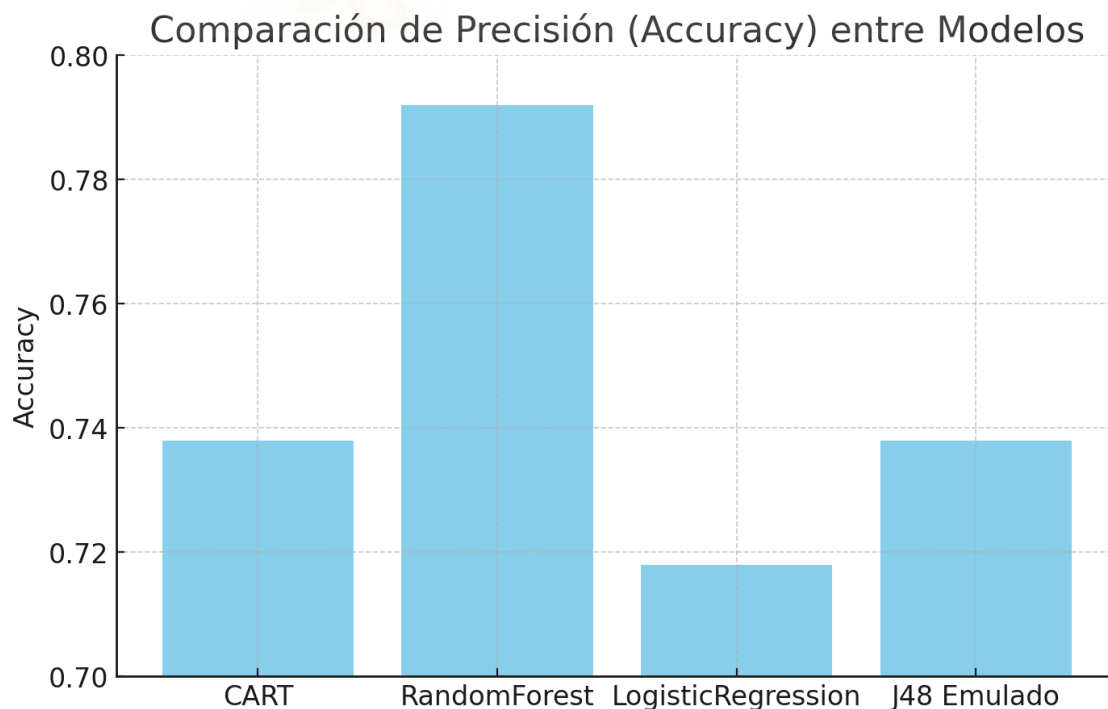


Figura 7.18: Comparación de Precisión entre los Modelos en el Entorno de IA.

En la Figura 7.18, se pueden observar las siguientes características en cuanto a la precisión de los modelos evaluados en el entorno apoyado por IA:

- **randomForest:** El modelo randomForest obtiene la mayor precisión entre todos los modelos, alcanzando un valor cercano al 0.8. Esto refleja un buen rendimiento en la clasificación de las instancias dentro del entorno de IA.
- **CART:** El modelo CART presenta una precisión de alrededor de 0.74, siendo inferior a randomForest, pero aún mostrando un desempeño aceptable en el entorno de IA.
- **LogisticRegression:** El modelo de Regresión Logística tiene la precisión más baja entre los modelos evaluados, con un valor ligeramente superior a 0.7, lo que indica que es menos eficaz en este entorno en comparación con otros modelos.
- **J48 Emulado:** El modelo J48 Emulado presenta una precisión cercana a 0.75, lo que lo sitúa por debajo de randomForest, pero ligeramente superior a CART y Regresión Logística.

Las diferencias observadas en la precisión entre los modelos sugieren que randomForest es el modelo que mejor se adapta al entorno apoyado por IA, seguido por J48 Emulado. Por otro lado, la Regresión Logística parece ser el modelo menos efectivo en este entorno. Estas variaciones en el rendimiento deben ser tenidas en cuenta al seleccionar un modelo para su implementación práctica.

7.5.3. Comparación entre el entorno de IA y el entorno local

Al comparar los resultados del entorno de IA con el entorno local, se observa que el modelo randomForest sigue siendo el más preciso en ambos entornos en cuanto a precisión se refiere, aunque en el entorno de IA obtiene una ligera ventaja con una precisión cercana al 0.8, mientras que en el entorno local su precisión es de aproximadamente 0.7. Por otro lado, modelos como LogisticRegression muestran una caída más pronunciada en el entorno local, donde la precisión es inferior, lo que

indica que el entorno de IA ofrece mejores condiciones para este tipo de algoritmos. En general, los modelos tienden a desempeñarse mejor en el entorno de IA, aunque las diferencias no son drásticas.

7.6. Conclusiones

El análisis comparativo entre los modelos y los conjuntos de datos A y B ha permitido extraer varias conclusiones importantes que no solo destacan las diferencias en rendimiento entre los algoritmos evaluados, sino también sus implicaciones prácticas en el contexto educativo.

▪ Dataset A:

- El modelo **RandomForest** ha demostrado ser el más adecuado para este dataset, alcanzando una precisión significativamente mayor en comparación con otros modelos, como J48 y LogisticRegression. Esto se evidencia tanto en la matriz de confusión como en la curva ROC, que indican una mejor capacidad de clasificación y discriminación, lo que sugiere que RandomForest puede identificar patrones complejos en los datos con mayor eficacia.
- Sin embargo, es importante señalar que el número de falsos positivos y falsos negativos, aunque reducido, sigue siendo considerable, lo que indica que este modelo aún tiene margen de mejora en la clasificación de estudiantes en riesgo. En un contexto práctico, esto podría significar que, si bien RandomForest ayuda a identificar con mayor precisión a los estudiantes que necesitan apoyo, es posible que algunos casos de riesgo pasen desapercibidos, y que otros sean clasificados incorrectamente, lo cual podría derivar en intervenciones innecesarias.
- El uso del entorno apoyado por IA mejoró ligeramente la capacidad discriminativa de RandomForest, lo que sugiere que la implementación de soluciones de IA podría ser beneficiosa en entornos educativos que busquen optimizar sus sistemas de predicción académica.

■ **Dataset B:**

- En el dataset B, el modelo **LogisticRegression** se destacó por su capacidad discriminativa, a pesar de tener una precisión general más baja que en el dataset A. La curva ROC mostró que, para este dataset, LogisticRegression es más eficaz al distinguir entre estudiantes de diferentes categorías de rendimiento académico.
- A pesar de ello, el número de falsos positivos y falsos negativos sigue siendo elevado, lo que sugiere que, en la práctica, este modelo podría llevar a errores significativos en la identificación de estudiantes en riesgo de bajo rendimiento. Esto plantea un reto importante, ya que las instituciones educativas que utilicen LogisticRegression en este contexto podrían implementar intervenciones basadas en predicciones incorrectas, lo cual puede afectar negativamente la eficiencia de los recursos educativos.
- El hecho de que LogisticRegression funcione mejor en este dataset subraya la importancia de adaptar los modelos a las características específicas de los datos utilizados. Las implicaciones prácticas de esta observación refuerzan la necesidad de evaluar y seleccionar el modelo correcto según el perfil de los estudiantes y los datos que se estén analizando.

■ **Comparación entre modelos:**

- En general, la precisión de los modelos varía considerablemente entre los dos conjuntos de datos. Mientras que **RandomForest** sobresale en el dataset A, **LogisticRegression** se desempeña mejor en el dataset B en términos de capacidad discriminativa. Estas diferencias pueden estar relacionadas con la naturaleza de los datos, como la diversidad en las características estudiadas, lo que subraya la importancia de realizar una selección adecuada del modelo basado en las propiedades específicas de los datos.
- El entorno apoyado por IA mostró mejoras generalizadas en el rendimiento de los modelos, especialmente en términos de capacidad discriminativa y una ligera reducción en los falsos negativos. En un contexto práctico, esto sugiere que el uso de herramientas de IA no solo mejora la precisión

de los modelos, sino que también puede contribuir a intervenciones más precisas y oportunas en el entorno educativo. Esto tiene implicaciones directas para la toma de decisiones informadas y personalizadas en la gestión de estudiantes en riesgo.

- Sin embargo, a pesar de estas mejoras, tanto en el entorno local como en el entorno apoyado por IA, los resultados indican que ninguno de los modelos es completamente preciso. Esto resalta la necesidad de continuar investigando formas de mejorar la precisión y minimizar los errores de clasificación, especialmente en aplicaciones prácticas donde las decisiones basadas en estos modelos pueden tener un impacto directo en los estudiantes.

■ **Conclusión global:**

- En términos generales, **RandomForest** se perfila como el modelo más adecuado para el dataset A, mientras que **LogisticRegression** es preferible para el dataset B debido a su mayor capacidad discriminativa en este último. La variabilidad en los resultados destaca la importancia de adaptar el modelo según las características específicas de los datos.
- En un contexto práctico, estos hallazgos sugieren que se deben considerar cuidadosamente las particularidades de sus datos antes de implementar un modelo predictivo. Además, las mejoras observadas con el uso de entornos apoyados por IA refuerzan la necesidad de incorporar soluciones tecnológicas avanzadas para mejorar la precisión y eficiencia en la predicción del rendimiento académico.

Por último, este trabajo también aborda la importancia de desarrollar herramientas tecnológicas de manera ética. La recolección y análisis de grandes volúmenes de datos plantea retos significativos en cuanto a la privacidad y la protección de datos. En este sentido, se ha seguido un enfoque que garantiza el cumplimiento de las normativas de protección de datos, como el Reglamento General de Protección de Datos (GDPR).

La predicción del rendimiento académico mediante ML no solo mejora la capacidad

de respuesta del sistema educativo, sino que sienta las bases para el desarrollo de futuras herramientas que impulsen la calidad de la enseñanza. Este TFG no solo ha mostrado un ejemplo de la evaluación y desarrollo de algoritmos, sino también subraya la necesidad de un uso responsable y ético de estas tecnologías.

7.7. Recomendaciones para futuras investigaciones

Esta investigación ha demostrado la viabilidad del uso de algoritmos de ML para la predicción del rendimiento académico, sin embargo, este es solo el primer paso, ya que se puede mejorar de distintas formas. A continuación, se detallan algunas recomendaciones específicas:

Nuevos conjuntos de datos: Los resultados podrían validarse o mejorarse con el uso de datasets adicionales que representen distintos contextos educativos. Un ejemplo es el *dataset KDD Cup 2010 Educational Data Mining Challenge*, que contiene información sobre los clics de los estudiantes en plataformas de aprendizaje en línea, lo que permitiría analizar su comportamiento en e-learning y compararlo con entornos presenciales.

Otro dataset útil es el *ASSISTments Student Performance Dataset*, que registra respuestas a ejercicios matemáticos, permitiendo probar la generalización de los modelos predictivos en disciplinas específicas como las matemáticas.

Futuros estudios también podrían usar datasets multilingües y multiculturales, como los de la Organización para la Cooperación y el Desarrollo Económico (OECD) (*PISA*), para investigar si los algoritmos entrenados en un contexto cultural pueden transferirse a otros.

Optimización de hiperparámetros: Se usaron hiperparámetros estándar en esta investigación, pero futuras optimizaciones podrían mejorar los resultados. Se recomienda una búsqueda exhaustiva de hiperparámetros utilizando técnicas como *grid search* o *random search*, ajustando parámetros como el número de árboles en *RandomForest*, la profundidad máxima en *CART*, o el valor de regularización en la *regresión logística*.

Otra técnica más eficiente es la *optimización bayesiana*, que ajusta los hiperparámetros basándose en el rendimiento de los modelos anteriores, reduciendo la necesidad de búsquedas exhaustivas.

También se podría explorar la *validación cruzada en múltiples niveles* (*nested cross-validation*), especialmente en escenarios con pocos datos, para evitar el sobreajuste y mejorar la generalización del modelo.

Técnicas avanzadas de ML: El uso de técnicas más avanzadas como las *redes neuronales profundas* (*deep learning*) podría mejorar los resultados en problemas con grandes volúmenes de datos y alta dimensionalidad. Estas técnicas podrían aplicarse a datasets más complejos que incluyan datos multimedia o texto no estructurado, como el análisis de sentimientos en comentarios de estudiantes.

Además, el uso de *modelos híbridos* que combinen enfoques supervisados y no supervisados puede ser útil. Por ejemplo, *K-means* podría segmentar a los estudiantes en grupos antes de aplicar un modelo predictivo supervisado, mejorando la precisión en grupos específicos.

Mitigación de sesgos: Futuros estudios deberían centrarse en reducir los sesgos en los modelos predictivos. Se recomienda realizar auditorías periódicas para garantizar que las predicciones sean justas y no estén influenciadas por factores sensibles como el género o la etnia. Técnicas como *Equalized Odds* o *Fairness through Awareness* pueden ayudar a evitar desigualdades en la asignación de recursos o intervenciones educativas.

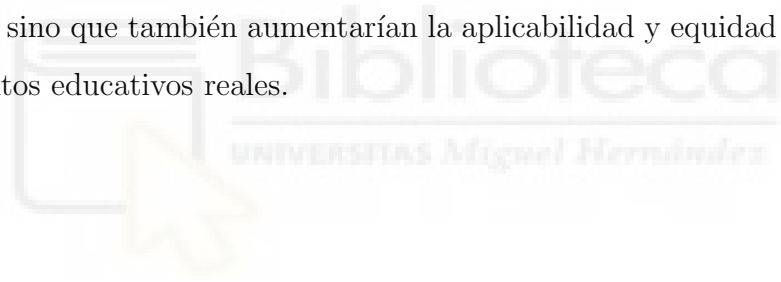
Además, sería útil investigar la creación de datasets sintéticos balanceados para evaluar mejor las estrategias de mitigación de sesgos.

Aplicación de técnicas de interpretabilidad: Dado que las predicciones se utilizarán para tomar decisiones importantes en la educación, es fundamental garantizar que los modelos sean comprensibles para los docentes. Futuros estudios podrían aplicar técnicas avanzadas de interpretabilidad, como *LIME* o *SHAP*, para hacer que los modelos predictivos sean más transparentes. Estas técnicas permiten desglosar las predicciones y mostrar los factores más importantes que influyen en el rendimiento

de los estudiantes, facilitando la confianza y la adopción de los modelos en entornos educativos.

Extensión del análisis longitudinal: Los modelos predictivos podrían mejorar al aplicar un enfoque longitudinal que estudie los patrones de los estudiantes a lo largo de varios años o etapas educativas. En lugar de enfocarse en un solo curso o semestre, futuros estudios podrían utilizar modelos secuenciales, como los *modelos de Markov ocultos (HMM)* o las *redes neuronales recurrentes (RNN)*, para analizar cómo las interacciones y comportamientos a lo largo del tiempo influyen en el rendimiento académico.

En resumen, futuras investigaciones deberían enfocarse en la diversificación de los datasets, la optimización avanzada de hiperparámetros, la aplicación de técnicas más sofisticadas de ML, la mitigación del sesgo algorítmico, y la mejora de la interpretabilidad de los modelos. Estos pasos no solo mejorarían la precisión de las predicciones, sino que también aumentarían la aplicabilidad y equidad de los modelos en contextos educativos reales.



BIBLIOGRAFÍA

- [1] C. R. y Sebastián Ventura, «Data mining in education,» *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 3, n.º 1, 2013.
- [2] P. V. y Axel von dem Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*. Springer, 2017.
- [3] S. B. y Andrew D Selbst, «Big data's disparate impact,» *California Law Review*, vol. 104, n.º 3, 2016.
- [4] C. D. y John Richards, «Opportunities and challenges of digital learning,» *The Journal of Educational Research*, vol. 106, n.º 4, 2014.
- [5] T. H. y R. T. Jerome Friedman, *Machine learning and education: A survey*. Springer, 2017.
- [6] L. Johnson et al., «Educational data mining and learning analytics: Potentials and challenges,» *Journal of Educational Technology & Society*, vol. 17, n.º 4, págs. 1-2, 2014.
- [7] C. Romero y S. Ventura, «Data mining in education,» *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 3, n.º 1, págs. 12-27, 2013.
- [8] R. S. Baker y K. Yacef, «Challenges for the future of educational data mining: The Baker Learning Analytics Prizes,» *Journal of Educational Data Mining*, vol. 11, n.º 1, págs. 1-17, 2019.
- [9] D. H. Wolpert y W. G. Macready, *No Free Lunch Theorems for Optimization*. IEEE Transactions on Evolutionary Computation, 1997, vol. 1, págs. 67-82.

- [10] V. Tinto, *Research and practice of student retention: What next?* Journal of College Student Retention: Research, Theory & Practice, 2006.
- [11] I. H. W. y Eibe Frank y Mark A. Hall y Christopher J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2016.
- [12] C. R. y Sebastián Ventura, «Educational data mining: a review of the state of the art,» *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 40, n.º 6, 2010.
- [13] S. Huang y N. Fang, «Data mining and analytics in higher education,» *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 4, n.º 2, 2014.
- [14] Q. Y. y DiGangi, «Intelligent tutoring systems: A comprehensive review,» *Journal of Educational Computing Research*, vol. 56, n.º 1, 2017.
- [15] J. R. y Murphy, *Personalized learning: A guide for engaging students with technology*. John Wiley & Sons, 2017.
- [16] R. Ferguson, «Learning analytics: Challenges and opportunities,» *Handbook of learning analytics*, 2016.
- [17] M. M. M. y Salleh, «Sentiment analysis for education service,» *World Academy of Science, Engineering and Technology*, vol. 69, 2012.
- [18] D. W. y Yang, «Sentiment analysis in MOOC discussion forums: What does it tell us?» *Proceedings of the 7th International Conference on Educational Data Mining*, 2014.
- [19] Y. L. y Yoshua Bengio y Geoffrey Hinton, «Deep Learning,» *Nature*, vol. 521, n.º 7553, 2015.

- [20] W. S. M. y Walter Pitts, «A Logical Calculus of the Ideas Immanent in Nervous Activity,» *The Bulletin of Mathematical Biophysics*, vol. 5, n.º 4, págs. 115-133, 1943.
- [21] V. Cortes Corinna y Vapnik, «Support-vector networks,» *Machine Learning*, vol. 20, n.º 3, 1995.
- [22] Future Space. «Machine Learning: los orígenes y la evolución.» Último acceso: 01 Sep 2024. (2024), dirección: <https://www.futurespace.es/machine-learning-los-origenes-y-la-evolucion/>.
- [23] M. A. y otros, «TensorFlow: A System for Large-Scale Machine Learning,» en *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, 2016.
- [24] A. P. y otros, «PyTorch: An Imperative Style, High-Performance Deep Learning Library,» en *Advances in Neural Information Processing Systems*, vol. 32, 2019.
- [25] I. G. y Yoshua Bengio y Aaron Courville, *Deep Learning*. MIT Press, 2016.
- [26] M. T. R. y Sameer Singh y Carlos Guestrin, «Why Should I Trust You?: Explaining the Predictions of Any Classifier,» en *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016.
- [27] S. M. L. y Su-In Lee, «A Unified Approach to Interpreting Model Predictions,» *Advances in Neural Information Processing Systems*, vol. 30, 2017.
- [28] S. B. y Andrew D. Selbst, «Big Data's Disparate Impact,» en *California Law Review*, vol. 104, UC Berkeley School of Law, 2016, págs. 671-732.

- [29] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- [30] J. F. y Trevor Hastie y Robert Tibshirani, «The Elements of Statistical Learning,» *Springer Series in Statistics*, 2001.
- [31] T. H. y Robert Tibshirani y Jerome Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2009.
- [32] G. J. y Daniela Witten y Trevor Hastie y Robert Tibshirani, *An Introduction to Statistical Learning*. Springer, 2013.
- [33] R. S. S. y Andrew G. Barto, *Reinforcement Learning: An Introduction*. MIT Press, 2018.
- [34] V. M. y Koray Kavukcuoglu y David Silver y Andrei A. Rusu y Joel Veness y Marc G. Bellemare y Alex Graves y Martin Riedmiller y Andreas K. Fidjeland y Georg Ostrovski y otros, «Human-level control through deep reinforcement learning,» *Nature*, vol. 518, n.º 7540, 2015.
- [35] X. H. y Lizi Liao y Hanwang Zhang y Li Nie y Xia Hu y Tat-Seng Chua, «Neural Collaborative Filtering,» *Proceedings of the 26th International Conference on World Wide Web*, 2017.
- [36] D. C. M. y Elizabeth A. Peck y G. Geoffrey Vining, *Introduction to Linear Regression Analysis*. John Wiley Sons, 2021.
- [37] S. L. y R. X. Hosmer David W., *Applied Logistic Regression*. John Wiley & Sons, 2013.
- [38] R. O. y C. S. Leo Breiman Jerome Friedman, *Classification and Regression Trees*. Wadsworth Brooks, 1984.

- [39] L. Breiman, «Random Forests,» *Machine Learning*, vol. 45, n.º 1, 2001.
- [40] N. S. Altman, «An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression,» *The American Statistician*, vol. 46, n.º 3, 1992.
- [41] I. G. y Andre Elisseeff, «An Introduction to Variable and Feature Selection,» *Journal of Machine Learning Research*, vol. 3, 2003.
- [42] S. W. y Kim Esbensen y Paul Geladi, «Principal Component Analysis,» *Chemometrics and Intelligent Laboratory Systems*, vol. 2, n.º 1-3, 1987.
- [43] R. J. L. y Donald B. Rubin, *Statistical Analysis with Missing Data*. John Wiley & Sons, 2019.
- [44] A. M. y Sarah Guido, *Introduction to Machine Learning with Python: A Guide for Data Scientists*. O'Reilly Media, 2016.
- [45] P. J. Rousseeuw, «A Graphical Aid to the Interpretation and Validation of Cluster Analysis,» *Journal of Computational and Applied Mathematics*, vol. 20, 1993.
- [46] R. A. J. y Dean W. Wichern, *Applied Multivariate Statistical Analysis*. Pearson, 2012.
- [47] T. Fawcett, «An introduction to ROC analysis,» *Pattern Recognition Letters*, vol. 27, n.º 8, 2006.
- [48] S. A. y Alain Celisse, «A Survey of Cross-Validation Procedures for Model Selection,» *Statistics Surveys*, vol. 4, 2010.
- [49] R. Tibshirani, «Regression Shrinkage and Selection via the Lasso,» *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, n.º 1, 1996.

- [50] S. Raschka, *Python Machine Learning*. Packt Publishing Ltd, 2015.
- [51] A. E. H. y Robert W. Kennard, «Ridge Regression: Biased Estimation for Nonorthogonal Problems,» *Technometrics*, vol. 12, n.º 1, 1970.
- [52] H. Z. y Trevor Hastie, «Regularization and Variable Selection via the Elastic Net,» *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 67, n.º 2, 2005.
- [53] N. S. y Geoffrey Hinton y Krizhevsky, «Dropout: A Simple Way to Prevent Neural Networks from Overfitting,» *Journal of Machine Learning Research*, vol. 15, n.º 56, 2014.
- [54] L. Prechelt, «Early Stopping - But When?» En *Neural Networks: Tricks of the Trade*, Springer, 1998.
- [55] J. B. y Rémi Bardenet y Yoshua Bengio y Balázs Kégl, «Algorithms for Hyper-Parameter Optimization,» *Advances in Neural Information Processing Systems*, vol. 24, 2011.
- [56] N. P. J. y otros, «In-Datacenter Performance Analysis of a Tensor Processing Unit,» *Proceedings of the 44th Annual International Symposium on Computer Architecture*, 2017.
- [57] M. Z. y otros, «Spark: Cluster Computing with Working Sets,» en *2nd USENIX Workshop on Hot Topics in Cloud Computing*, 2010.
- [58] T. K. y otros, «Jupyter Notebooks—a Publishing Format for Reproducible Computational Workflows,» en *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, 2016.

- [59] M. Z. y otros, «Accelerating the Machine Learning Lifecycle with MLflow,» en *Data Engineering*, 2018.
- [60] F. H. y Lars Kotthoff y Joaquin Vanschoren, «Automated Machine Learning: Methods, Systems, Challenges,» *Springer Nature*, 2019.
- [61] M. V. y otros, «Infrastructure Cost Comparison of Running Web Applications in the Cloud Using AWS Lambda and Monolithic and Microservice Architectures,» en *2018 IEEE/ACM 11th International Conference on Utility and Cloud Computing (UCC)*, 2018.
- [62] H. Yu y S. DiGangi, «A Decision Tree Based Predictive Model for Student Success in Online Learning,» *Journal of Educational Technology*, vol. 4, n.º 3, págs. 45-50, 2017.
- [63] M. Wen, D. Yang y C. P. Rose, «Sentiment analysis in education: Students' engagement in online discussions,» *Educational Data Mining*, vol. 7, n.º 3, págs. 30-40, 2014.
- [64] E. Union, *General Data Protection Regulation (GDPR)*. Official Journal of the European Union, 2018. dirección: <https://eur-lex.europa.eu/eli/reg/2016/679/oj>.
- [65] M. Feldman, S. A. Friedler, J. Moeller, C. Scheidegger y S. Venkatasubramanian, «Certifying and removing disparate impact,» en *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015, págs. 259-268.
- [66] S. Wachter, B. Mittelstadt y L. Floridi, «Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation,» *International Data Privacy Law*, vol. 7, n.º 2, págs. 76-99, 2017.

- [67] R. Binns, «Fairness in Machine Learning: Lessons from Political Philosophy,» *Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency*, págs. 149-159, 2018.
- [68] C. Rudin, «Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead,» *Nature Machine Intelligence*, vol. 1, n.º 5, págs. 206-215, 2019.
- [69] U. F. y Gregory Piatetsky-Shapiro y Padhraic Smyth, «From Data Mining to Knowledge Discovery in Databases,» *AI Magazine*, vol. 17, n.º 3, 1996.
- [70] S.-H. L. y Paul-Huang Chu y Pei-Yuan Hsiao, «Data Mining Techniques and Applications – A Decade Review from 2000 to 2011,» *Expert Systems with Applications*, vol. 39, n.º 12, 2012.
- [71] J. R. Quinlan, «Induction of Decision Trees,» en *Machine Learning*, vol. 1, 1986.
- [72] A. K. J. y M. Narasimha Murty y Patrick J. Flynn, «Data Clustering: A Review,» *ACM Computing Surveys*, vol. 31, n.º 3, 1999.
- [73] R. A. y Ramakrishnan Srikant, «Fast Algorithms for Mining Association Rules,» en *Proceedings of the 20th International Conference on Very Large Data Bases (VLDB)*, 1994.
- [74] V. C. y Arindam Banerjee y Vipin Kumar, «Anomaly Detection: A Survey,» *ACM Computing Surveys*, vol. 41, n.º 3, 2009.
- [75] I. T. J. y Jorge Cadima, *Principal Component Analysis*. Springer, 2016.
- [76] C. D. M. y Prabhakar Raghavan y Hinrich Schütze, *Introduction to Information Retrieval*. Cambridge University Press, 2014.

- [77] C. X. L. y Chenghui Li, «Data Mining for Direct Marketing: Problems and Solutions,» en *Proceedings of the 4th International Conference on Knowledge Discovery and Data Mining (KDD)*, 1998.
- [78] R. J. B. y David J. Hand, «Statistical Fraud Detection: A Review,» *Statistical Science*, vol. 17, n.º 3, 2002.
- [79] M. M. Gaber, *Data Mining for Health Informatics*. Springer, 2010.
- [80] C.-F. T. y Chih-Wei Lai y Chao, «Big Data Analytics: A Survey,» *Journal of Big Data*, vol. 2, n.º 1, 2014.
- [81] J. H. y Tahar M. Kechadi, «Telecom Customer Segmentation Based on Cluster Analysis,» *Procedia Computer Science*, vol. 60, 2015.
- [82] R. S. B. y Paul S. Inventado, «Educational Data Mining and Learning Analytics,» *Springer*, 2014.
- [83] B. M. y Robert Cooley y Jaideep Srivastava, «Web Usage Mining and Personalization,» en *Communications of the ACM*, vol. 43, 2001.
- [84] H. C. y otros, «Data Mining for the Masses,» *Communications of the ACM*, vol. 47, n.º 11, 2004.
- [85] A. M. Turing, «Computing machinery and intelligence,» *Mind*, vol. 59, n.º 236, págs. 433-460, 1950.
- [86] D. Crevier, *AI: The Tumultuous History of the Search for Artificial Intelligence*. Basic Books, 1993.
- [87] S. Russell, *Artificial Intelligence: A Modern Approach*, 3.ª ed. Prentice Hall, 2009.

- [88] N. J. Nilsson, *Artificial Intelligence: A New Synthesis*. Morgan Kaufmann, 1998.
- [89] N. Bostrom, *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, 2014.
- [90] OECD, «OECD Recommendation of the Council on Artificial Intelligence,» 2019.
- [91] J. M. y Marvin L. Minsky y Nathaniel Rochester y Claude E. Shannon, «A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence,» en *Dartmouth Conference on Artificial Intelligence*, Dartmouth College, 1956.
- [92] A. N. y Shaw, «The Logic Theory Machine: A Complex Information Processing System,» *IRE Transactions on Information Theory*, vol. 2, n.º 3, págs. 61-79, 1956.
- [93] P. J. Werbos, «Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences,» en *PhD Thesis*, Harvard University, 1974.
- [94] R. S. S. y Andrew G. Barto, *Reinforcement Learning: An Introduction*, 2.ª ed. MIT Press, 2018.
- [95] E. M. B. y Timnit Gebru y McMillan-Major, «On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?» *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, págs. 610-623, 2021.
- [96] E. Union, *Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*, 2021. direc-

- ción: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>.
- [97] E. Union, *General Data Protection Regulation*, 2016. dirección: <https://eur-lex.europa.eu/eli/reg/2016/679/oj>.
- [98] C. Rudin, «Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead,» *Nature Machine Intelligence*, vol. 1, págs. 206-215, 2019.
- [99] E. Commission, *White Paper on Artificial Intelligence: A European approach to excellence and trust*, 2020. dirección: https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en.
- [100] L. F. y C. Cowl, «AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations,» *Minds and Machines*, vol. 28, n.º 4, págs. 689-707, 2018.
- [101] R. Ferguson, «Learning Analytics: Drivers, Developments and Challenges,» *International Journal of Technology Enhanced Learning*, vol. 4, n.º 5-6, págs. 304-317, 2012.
- [102] J. P. C. y Diana G. Oblinger, «Early Alert: Using Learning Analytics to Predict Academic Success,» en *Proceedings of the EDUCAUSE Annual Conference*, 2007.
- [103] J. Piaget, *The Origins of Intelligence in Children*. International Universities Press, 1952.

- [104] G. Siemens, «Connectivism: A Learning Theory for the Digital Age,» *International Journal of Instructional Technology and Distance Learning*, vol. 2, n.º 1, 2005.
- [105] B. J. Zimmerman, «Becoming a Self-Regulated Learner: An Overview,» *Theory Into Practice*, vol. 41, n.º 2, 2002.
- [106] C. R. y Sebastián Ventura, «Educational Data Mining: A Review of the State of the Art,» *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 40, n.º 6, 2010.
- [107] D. G. y otros, «Learning Analytics: Challenges and Opportunities,» en *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*, 2014.
- [108] S. Dawson, «“Seeing” the Learning Community: An Exploration of the Development of a Resource for Monitoring Online Student Networking,» *British Journal of Educational Technology*, vol. 41, n.º 5, 2010.
- [109] A. P. y George Siemens, «Personalised Learning Analytics: A Conceptual Framework and Pilot Study,» *Journal of Learning Analytics*, vol. 4, n.º 1, 2017.
- [110] G. S. y Phil Long, «Learning and Knowledge Analytics,» en *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, 2011.
- [111] G. S. y Ryan SJD Baker, *Handbook of Learning Analytics*. Society for Learning Analytics Research (SoLAR), 2013.
- [112] P. P. y Sharon Slade, «The Ethics of Learning Analytics: The ‘CREEPER’ Framework,» *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, págs. 153-155, 2013.

- [113] A. P. y George Siemens, «Ethical and Privacy Principles for Learning Analytics,» en *British Journal of Educational Technology*, vol. 45, 2012.
- [114] B. Williamson, «Big Data in Education: The Digital Future of Learning, Policy and Practice,» *Sage Publications*, 2017.
- [115] K. Bruns, «Visual Studio Code: End-to-End Editing and Debugging Tools for Web Developers,» *Web Developer Journal*, 2020.
- [116] A. G. y otros, «Kaggle: Your Home for Data Science,» *Kaggle Blog*, 2010.
- [117] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell et al., «Language Models are Few-Shot Learners,» *Advances in Neural Information Processing Systems*, vol. 33, págs. 1877-1901, 2020.
- [118] T. O. University. «Open University Learning Analytics Dataset.» (2024), dirección: https://analyse.kmi.open.ac.uk/open_dataset.