

Mejoramiento académico en instituciones educativas aplicando algoritmos de Machine Learning.

Davinson Anaya Sierra, ✉ davison.anaya201@tau.usbmed.edu.co
Santiago Munera Arango, ✉ santiago.munera201@tau.usbmed.edu.co
Tatiana Salazar Bedoya, ✉ tatiana.salazar201@tau.usbmed.edu.co

Trabajo de Grado presentado para optar al título de Ingeniero de Datos y Software

Asesor: Ph.D. Andrés Arbey Ochoa Rojas



Universidad de San Buenaventura Colombia
Facultad de Ingenierías
Ingeniería de Datos y Software
Medellín
2023

Citar/How to cite

[1]

Referencia Bibtex

```
@masterthesis{Munera Arango-Salazar Bedoya-Anaya  
Sierra2023,  
author = {Davinson Anaya Sierra and Santiago Munera Arango  
and Tatiana Salazar Bedoya},  
title = {Mejoramiento académico en instituciones educativas  
aplicando algoritmos de Machine Learning.},  
school = {Universidad de San Buenaventura},  
type = {Tesis de Pregrado},  
year = {2023}  
}
```

Referencia

Estilo IEEE 2014

Davinson Anaya Sierra and Santiago Munera Arango and Tatiana Salazar Bedoya, "Mejoramiento académico en instituciones educativas aplicando algoritmos de Machine Learning.", Tesis de Pregrado, Ingeniería de Datos y Software, Universidad de San Buenaventura, Facultad de Ingenierías, 2023

En convenio con la Universidad (nombre y logo pequeño de la institución).

Seleccione posgrado USB Colombia, Cohorte X.

Grupo de Investigación (SIGLA).

Línea de investigación en Acústica y Procesamiento de señales.

Bibliotecas Universidad de San Buenaventura



Biblioteca Digital (Repositorio)
<http://bibliotecadigital.usb.edu.co>

- Biblioteca Fray Alberto Montealegre OFM - Bogotá.
- Biblioteca Fray Arturo Calle Restrepo OFM - Medellín, Bello, Armenia, Ibagué.
- Departamento de Biblioteca - Cali.
- Biblioteca Central Fray Antonio de Marchena – Cartagena.

Universidad de San Buenaventura Colombia

Universidad de San Buenaventura Colombia - <http://www.usb.edu.co/>

Bogotá - <http://www.usbbog.edu.co>

Medellín - <http://www.usbmed.edu.co>

Cali - <http://www.usbcali.edu.co>

Cartagena - <http://www.usbctg.edu.co>

Editorial Bonaventuriana - <http://www.editorialbonaventuriana.usb.edu.co/>

Revistas - <http://revistas.usb.edu.co/>

Dedicatorias

Texto de dedicatoria centrado.

Agradecimientos

Texto de agradecimiento centrado.

TABLA DE CONTENIDOS

RESUMEN	7
ABSTRACT	7
I. INTRODUCCIÓN	8
A. Ejemplos de citación	9
II. ANTECEDENTES	10
A. Antecedentes Históricos	10
1. Rendimiento academico	10
2. Machine Learning	10
3. Clasificación/predicción	11
4. Minería de datos	11
B. Antecedentes de la Investigación	11
1. "Predicción del rendimiento académico como indicador de éxito/fracaso de los estudiantes de ingeniería, mediante aprendizaje automático"	11
2. "Análisis comparativo de algoritmos de aprendizaje automático supervisado para construir un modelo predictivo para evaluar el rendimiento de los estudiantes"	11
3. "Machine Learning para predecir el rendimiento académico de los estudiantes universitarios"	12
4. "Deserción universitaria: Evaluación de diferentes algoritmos de Machine Learning para su predicción"	12
5. "Modelos predictivos de riesgo académico en carreras de computación con minería de datos educativos"	12
6. "Aplicación de técnicas de minería de datos para la evaluación del rendimiento académico y la deserción estudiantil"	13
III. PLANTEAMIENTO DEL PROBLEMA	14
A. Pregunta de investigación	14
B. Hipótesis	14
IV. JUSTIFICACIÓN	15
V. OBJETIVOS	16
A. OBJETIVOS GENERAL	16
B. OBJETIVOS ESPECÍFICOS	16
VI. MARCO TEÓRICO	17
A. Rendimiento académico	17
B. Machine Learning (ML)	17
C. Tipos de algoritmos de ML	18
1. Aprendizaje supervisado	18
2. Aprendizaje no supervisado	20
3. Aprendizaje por refuerzo	21
D. Estadística	21
1. Error Cuadrático Medio (MSE)	21
2. Error Absoluto Medio (MAE)	22
3. Raíz del Error Cuadrático Medio (RMSE)	22

VII. METODOLOGÍA	23
A. Descripción del dataset.	23
1. Atributos Generales.	23
2. Atributos Educativos.	23
3. Atributos Sociales.	24
4. Atributos relacionados con el rendimiento escolar.	24
B. Metodología	24
1. Determinar los modelos predictivos	24
2. Evaluar los modelos predictivos	24
3. Establecer el mejor modelo	25
4. Ejecutar algoritmos de aprendizaje	25
5. Análisis de datos obtenidos	25
VIII. Análisis exploratorio de datos	26
IX. CONCLUSIONES	33
X. RECOMENDACIONES	34
REFERENCIAS	35
Apéndice A: Ejemplo de Anexo: Datos	35
Apéndice B: Otro Apéndice	35

LISTA DE TABLAS

LISTA DE FIGURAS

RESUMEN

El resumen permite identificar la esencia del escrito, mencionando brevemente el objetivo y la metodología, así como los resultados y las conclusiones (mínimo 150, máximo 250 palabras).

Palabras clave: Palabra clave1, Palabra clave2, Palabra clave3.

ABSTRACT

El abstract es el mismo resumen pero en idioma inglés. Conserva la misma extensión o aproximada, es decir, mínimo 150 y máximo 250 palabras.

Keywords: Scientific article, Review article, Research, Citation styles.

I. INTRODUCCIÓN

En la actualidad los avances tecnológicos actuales y el mayor uso de sistemas y herramientas digitales han permitido a organizaciones de diversos sectores aplicar técnicas de análisis de datos y aprendizaje automático (ML) para aprovechar al máximo la información de que disponen y facilitar la toma de decisiones. Estos avances tecnológicos también se han aplicado al sector educativo, permitiendo una mayor interacción entre profesores, alumnos y recursos de aprendizaje dentro de las instituciones educativas. Esta mayor interacción ha llevado a una generación masiva de datos relacionados con las interacciones entre profesores y estudiantes, lo que presenta nuevos desafíos en cuanto al manejo de la información y los beneficios que se pueden obtener a partir de ella. Por lo tanto, las instituciones educativas necesitan transformar y analizar los datos que recopilan para medir, mejorar y adaptar sus prácticas y contenidos educativos a través de la analítica de datos. La analítica de datos en el ámbito educativo implica el uso de algoritmos y técnicas de ML para analizar y comprender mejor los datos estudiantiles y tomar decisiones más precisas y efectivas. La aplicación de algoritmos predictivos mediante técnicas de aprendizaje automático puede ser una herramienta útil para mejorar el rendimiento académico en los centros educativos. En la actualidad, las instituciones educativas disponen de una gran cantidad de información sobre sus alumnos, pero no cuentan con un sistema eficaz para analizar y evaluar estos datos con el fin de obtener informes precisos y tomar decisiones eficaces. Esta investigación tiene como objetivo evaluar y analizar datos estudiantiles de instituciones educativas utilizando algoritmos estadísticos y predictivos para mejorar la comprensión de los entornos de enseñanza, personalizar el proceso de aprendizaje de los estudiantes, brindar retroalimentación en tiempo real y predecir el rendimiento escolar. La utilización de técnicas de ML permite a las instituciones educativas tener un mejor entendimiento de los patrones de comportamiento de los estudiantes, lo que facilita la toma de decisiones informadas y efectivas. Además, la predicción de el rendimiento escolar ayuda a los educadores a identificar a tiempo a los estudiantes que necesitan ayuda adicional, lo que puede resultar en una mejora significativa en los pronósticos de evaluación. En la introducción se menciona claramente el para qué y el porqué del documento, es una combinación del planteamiento del problema, el objetivo principal, las preguntas de investigación y la justificación.

No utilice en el documento la primera persona en singular (yo realicé las encuestas) ni primera persona plural (realizamos las encuestas); utilice siempre la narración en tercera persona (se realizaron las encuestas, se publicaron resultados, se establecieron parámetros, etc.)

Por otro lado, los conectores son importantes en la estructura del texto, tener una buena variedad de estos enriquece la estructura y redacción del texto. Algunos ejemplos:

Sin embargo	Teniendo en cuenta
Puesto que	Entonces
Por consiguiente	Simultáneamente
Dado que	Posiblemente

En efecto	A saber
Ya que	De la misma forma
Ahora bien	En síntesis
En cambio	Así
En cuanto a	Para concluir
El siguiente punto es	Luego
Así pues	Resumiendo
Recapitulando	De igual manera
En conclusión	Al mismo tiempo
En pocas palabras	Probablemente
A continuación	Indiscutiblemente
Acto seguido	
Con motivo de	

A. Ejemplos de citación

L^AT_EX le permite referenciar diferente tipo de documento y seguir el estandar IEEE automáticamente. A continuación algunos ejemplos de referencias: un artículo *in proceeding* **elen2001ambisonics**, la siguiente es la citación de un *artículo* [1], Esta es de un *libro* **goossens93** y esta de una *página web* **knuthwebsite**. En la última parte de este texto se encuentra la tabla de referencia, allí puede observar el formato de impresión para cada tipo de documento. L^AT_EX ordena automáticamente las referencias por orden de aparición y las da un formato según el tipo de documento.

Para referencia dos fuentes este es el formato **hollerweger2013, elen2001ambisonics**, para mas fuentes consecutivas se vería así **greenwade93, elen2001ambisonics, hollerweger2013**. Nuestra plantilla usa la librería de biblatex para el manejo de referencias, mayor información sobre el tipo de referencias que se pueden citar y como citarlas pueden encontrarlas en la web: **Bibliography management with biblatex**.

II. ANTECEDENTES

A. Antecedentes Históricos

1) *Rendimiento académico*: Hay varias definiciones de rendimiento académico en el texto[1], una de las cuales define el rendimiento académico como la suma de varios factores complejos que actúan sobre las personas que esta en aprendizaje. Además, lo definen como el resultado de aprobar una materia, sin embargo, otros ven las calificaciones como una medida de los resultados de la enseñanza, más no una medida de calidad.

2) *Machine Learning*: Según el autor DENNIYE HINESTROZA[2] , el científico Alan Mathison Turing fue capaz de crear el llamado 'Test de Turing' en 1950, una prueba diseñada para medir la inteligencia de una computadora, al tratar de conversar con un humano de manera racional, teniendo conductas humanas. Desde entonces, en 1952, Arthut Samuel fue el hombre que escribió el primer algoritmo con competencias de aprender, una pieza de software con la capacidad de jugar ajedrez, mejorando cada vez en el juego, esto se debe a que, podía guardar información y distintas maneras de dar respuesta de acuerdo al avance del nivel. Posteriormente, en 1956, se origina en una conferencia científica de grandes personajes, el termino 'Inteligencia Artificial' el cual definiría el gran campo de la ciencia e informática.

De acuerdo con el artículo[3] , Machine Learning (ML) es la ciencia que posibilita el aprendizaje de las computadoras y preserva información en forma de observaciones e interacciones con el mundo real.

El ML[4] tiene dos tipos de aprendizajes: supervisado y no supervisado.

- El supervisado, son algoritmos previamente entrenados con características y etiquetas, con el fin de que en el futuro el algoritmo pueda realizar predicciones ya conociendo las características. Un ejemplo de este tipo de algoritmo: Naive Bayes, Árbol de decisión, Regresión Logística, K vecinos cercanos, etc[5].
- El no supervisado, son algoritmos a los cuales solo se le dan características y no etiquetas, con el fin de que agrupe los datos según las características. El algoritmo solo conoce los datos por como comparten algunas características y se asume que posiblemente pertenecen al mismo grupo. Cuando se obtiene un dato nuevo, este algoritmo lo agrupa de acuerdo al conjunto que tenga características más similares al dato inicial[5].

Los algoritmos de ML se pueden agrupar en 3 modelos: lineales, de árbol y redes neuronales, estos modelos nos ayudan en la identificación de patrones en datos masivos y en la elaboración de predicciones. Los campos de acción del machine learning son la medicina, la construcción, finanzas, robótica y la educación[2].

3) *Clasificación/predicción*: La clasificación/predicción de acuerdo a el autor[5], es una manera de analizar los datos que serán usados para la extracción de modelos que describirán los datos más importantes o para predecir el futuro comportamiento. Por otro lado, es el proceso de la construcción de un modelo que pueda describir el conjunto previo de clases, además, el modelo es empleado para generar la predicción de datos desconocidos, posteriormente, se estima la precisión del modelo. Por otro lado, varios algoritmos han sido propuestas para llevar acabo la predicción, la mayoría de algoritmos están habitados en memoria y se encargan de una pequeña cantidad de datos.

4) *Minería de datos*: El artículo[6] menciona que la minería de datos es el proceso de hallar anomalías, patrones y correlaciones en grandes conjuntos de datos para predecir resultados, esta requiere de operaciones que deben ser estudiadas por un estadístico alguien que entienda los conceptos y sepa cómo interpretarlos cuando ocurren cambios. Además, menciona que la minería de datos está comenzando a emplearse con bastante frecuencia debido a su conveniencia, transparencia y bajo costo de preservación de la información, y además, es una herramienta poderosa debido a su papel en el estudio de grandes bases de datos. El potencial de MD para trabajar con el tamaño de la base de datos a analizar es muy importante.

Según el autor[5], la minería de datos se utiliza en diversas industrias y diferentes funciones comerciales, tales como: empresas de telecomunicaciones, tarjetas de crédito y compañías de seguros, para la detección de fraudes, optimización de campañas de marketing, etc. Industrias de comercio, para diseñar y evaluar campañas de marketing, etc; En la medicina para predecir la efectividad de procedimientos quirúrgicos, exámenes médicos y medicamentos; bancos e instituciones financieras para mejorar el rendimiento económico.

B. Antecedentes de la Investigación

1) *"Predicción del rendimiento académico como indicador de éxito/fracaso de los estudiantes de ingeniería, mediante aprendizaje automático"*: Este trabajo corresponde a Leonardo E. Contreras, Héctor J. Fuentes y José I. Rodríguez [17] en el cual se plantearon analizar aquellas variables que influyeron en la predicción del rendimiento académico de estudiantes de ingeniería industrial de la Universidad Distrital de Colombia. La metodología implementada por los autores es el uso de algoritmos de clasificación por medio del lenguaje de programación Python, como arboles de decisión, K vecinos más cercanos, perceptrón entre otros. Estos algoritmos fueron comparados con el fin de conocer el modelo de mejor ajuste. De acuerdo a las variables utilizadas por los autores para dicho estudio, el modelo que mejor se ajustó a los datos con respecto al resto, fue el de perceptrón con una mejor exactitud. Como conclusión, las variables que más influyeron en este estudio fueron: edad, género, puntaje ICFES y valor de matrícula.

2) *"Análisis comparativo de algoritmos de aprendizaje automático supervisado para construir un modelo predictivo para evaluar el rendimiento de los estudiantes"*: Esta investigación corresponde a Inssaf El Guabassi, Marah Rim, Zakaria Bousalem y Aimad Qazda [18] donde tuvieron como objetivo

construir un modelo de clasificación con el fin de evaluar el rendimiento académico de los estudiantes. Por ende, como primer paso se aplicaron varios algoritmos de machine learning supervisados como ANCOVA, Regresión Logística, Regresión de Vecinos de Apoyo, Regresión Log-lineal, Regresión de Árbol de Decisión, Regresión de Formas Aleatorias y Regresión de Mínimos Cuadrados Parciales. El segundo paso, se realizaron comparaciones y evaluaciones de los algoritmos empleados para la creación del modelo basados en varias métricas de evaluación. Como último paso, se determinaron los factores de más relevancia que influyeron en el éxito o fracaso de los estudiantes. Como resultado, se obtuvo que el modelo de Regresión Log-lineal proporciona una mejor predicción, al igual que los factores de comportamiento influyen en el rendimiento académico.

3) *"Machine Learning para predecir el rendimiento académico de los estudiantes universitarios"*: García Dionisio y Jeancarlos Donato [19], elaboraron un modelo de Machine Learning con la finalidad de predecir el rendimiento académico, como metodología implementaron la KDD, además, herramientas como SPSS statistic y SPSS Modeler para la creación del modelo. Por otro lado, este estudio tuvo como fin determinar la precisión del machine learning en predecir el rendimiento de los estudiantes, con el objetivo de identificar la posibilidad de éxito o fracaso. El estudio tomó una población de estudio de 87 estudiantes, además, fue un estudio experimental. Este estudio concluyó, que los algoritmos usados como árbol de decisión, máquina de vectores y K-NN, quien tuvo mejores resultados con un 100 % de sensibilidad y especificidad fue la Máquina de vectores.

4) *"Deserción universitaria: Evaluación de diferentes algoritmos de Machine Learning para su predicción"*: Este trabajo corresponde a Valero Cajahuanca, Julio Elvis, Navarro Raymundo, Ángel Fernando, Larios Franco, Alfredo César, Julca Flores y Janett Deisy [20] quienes tuvieron como propósito establecer el algoritmo de Machine Learning de mejor ajuste para predecir la deserción universitaria. Estos autores se basaron en un estudio en Perú acerca de la deserción universitaria entre 2018 y 2021. De acuerdo con los autores, se usó una población de 652 estudiantes el 30 % fueron para entrenamiento y 70 % para prueba. Para el desarrollo de los modelos se utilizó el lenguaje de Python de Anaconda, con la ayuda de librerías. Como conclusión se obtuvo que el algoritmo K-Nearest-Neighbor con una precisión del 91 % para pronosticar la deserción de los estudiantes. Este modelo permite alertar a la institución para brindar ayudas académicas ya sea individual o grupal.

5) *"Modelos predictivos de riesgo académico en carreras de computación con minería de datos educativos"*: Enrique Ayala Franco, Rocío Edith López Martínez y Víctor Hugo Menéndez Domínguez [21] tuvieron como objetivo generar modelos predictivos de riesgos académicos, utilizando metodologías de la minería de datos educativa, que sirven como herramientas de detección temprana de riesgos académicos. En este estudio se tomaron datos sociodemográficos y resultados de exámenes de admisión. Como metodología emplearon las etapas del Proceso de Extracción de Conocimiento en Bases de Datos, se aplicaron las técnicas de clasificación para el análisis, obtención y validación de los modelos. Los resultados obtenidos muestran que el mejor modelo corresponde al algoritmo LMT con el 75.42 % y un 80.5 % para el área bajo la curva.

6) *"Aplicación de técnicas de minería de datos para la evaluación del rendimiento académico y la deserción estudiantil"*: Esta investigación realizada por Osvaldo M. Sposito, Osvaldo M. Sposito, Hugo L. Ryckeboer y Julio Bossero [22] quienes realizaron la aplicación de el proceso de descubrimiento de conocimiento sobre datos de estudiantes entre 2003 y 2008. Este proceso se realizó con el software MS SQL Server para el almacenamiento de los datos, el software SPSS para el preprocesamiento de los datos y Weka para encontrar un clasificador del rendimiento académico. El objetivo de los estudios es presentar un estudio empleando el proceso de DCDB por medio de clasificadores. Como resultados de la investigación, no se pudo encontrar un clasificador con alto grado de precisión, por ende, se buscará en estudios futuros una solución para la predicción del rendimiento de los estudiantes y su deserción.

III. PLANTEAMIENTO DEL PROBLEMA

El avance tecnológico y el aumento del uso de sistemas y herramientas digitales han permitido a empresas de diferentes industrias implementar técnicas de análisis de datos y machine learning con el objetivo de aprovechar al máximo la información que tienen y facilitar la toma de decisiones[7]. Asimismo, estos avances tecnológicos también se han trasladado al ámbito educativo, lo que ha permitido una mayor interacción entre profesores, estudiantes y recursos de aprendizaje dentro de las instituciones. Como resultado, las instituciones educativas generan una gran cantidad de datos relacionados con las interacciones entre profesores y estudiantes, lo que presenta nuevos desafíos en cuanto al manejo de la información y los beneficios que se pueden obtener a partir de ella. [8].

Por lo tanto, las instituciones educativas necesitan transformar y analizar los datos que recopilan para medir, mejorar y adaptar sus prácticas y contenidos educativos a través de la analítica de datos. Esto con el fin de mejorar el rendimiento académico de los estudiantes y, por ende, la calidad de la educación impartida por la institución y su modelo de enseñanza. Como investigador con experiencia en analítica de datos, propongo desarrollar un proyecto de rendimiento académico utilizando algoritmos de predicción y técnicas de machine learning para ayudar a las instituciones educativas en este proceso de mejora continua.

A. *Pregunta de investigación*

¿Qué algoritmo de aprendizaje automático predice el rendimiento académico de los estudiantes y qué modelo se ajusta mejor a los datos?

B. *Hipótesis*

Se busca implementar técnicas de machine learning que permitan analizar datos estudiantiles y predecir el desempeño académico de los estudiantes en instituciones educativas. La aplicación de algoritmos estadísticos permitirá obtener información relevante sobre el estado académico de los estudiantes, lo que ayudará a prevenir la deserción y a brindar retroalimentación oportuna para mejorar su proceso de aprendizaje. De esta manera, se espera lograr una mejora en la calidad de la educación y en el rendimiento académico de la institución.

IV. JUSTIFICACIÓN

El objetivo de esta investigación es evaluar y analizar datos estudiantiles [9] de instituciones educativas con el fin de generar informes y tomar decisiones más precisas y efectivas. Aunque todas las instituciones tienen una gran cantidad de información sobre sus estudiantes, actualmente no hay un sistema que evalúe, analice y genere informes sobre estos datos de manera que permita comprender y optimizar los entornos de enseñanza. Esta investigación propone el uso de algoritmos estadísticos y predictivos para revisar los métodos de enseñanza, personalizar el proceso de aprendizaje de los estudiantes, brindar retroalimentación en tiempo real y predecir la deserción y el rendimiento escolar para mejorar los pronósticos de evaluación.

La Universidad de San Buenaventura cuenta con un amplio conocimiento en el área de datos y técnicas de descubrimiento KDD (Knowledge Discovery in Databases), lo que le permite detectar patrones, comportamientos y tendencias en los datos. Los estudiantes de Ingeniería de Sistemas de la Universidad de San Buenaventura tienen acceso a estos conocimientos y pueden ejecutar proyectos relacionados. Además, la Universidad ha estado invirtiendo en infraestructura tecnológica para mejorar la calidad de sus procesos académicos.

V. OBJETIVOS

A. *OBJETIVOS GENERAL*

Analizar los datos estudiantiles con el uso de algoritmos predictivos y técnicas de Machine Learning para apoyar el mejoramiento académico de las instituciones educativas.

B. *OBJETIVOS ESPECÍFICOS*

- Obtener datos de muestra disponibles de las instituciones educativas y evaluar su calidad y relevancia para el análisis de rendimiento académico de los estudiantes.
- Investigar y seleccionar los algoritmos predictivos más adecuados dentro del conjunto de técnicas de Machine Learning, considerando su efectividad y aplicabilidad a los datos disponibles.
- Realizar una evaluación rigurosa de los modelos predictivos mediante técnicas de validación cruzada y otras herramientas de evaluación de desempeño.
- Realizar comparaciones entre los resultados obtenidos de diferentes modelos para determinar el mejor modelo para la implementación del análisis de datos.
- Implementar el modelo seleccionado para el análisis de los datos de rendimiento académico de los estudiantes y establecer los procedimientos para la interpretación de los resultados.
- Interpretar los resultados del análisis de datos, utilizando técnicas de aprendizaje de máquina y otros métodos estadísticos, para obtener información valiosa y significativa sobre el comportamiento académico de los estudiantes

VI. MARCO TEÓRICO

A. *Rendimiento académico*

El rendimiento académico se refiere al nivel de logro académico que alcanzan los estudiantes durante el proceso de aprendizaje y se mide mediante una serie de indicadores, como las notas, las calificaciones en los exámenes y la finalización de los cursos. El rendimiento académico es importante porque es un indicador clave del éxito de los estudiantes y puede repercutir en sus carreras educativas y profesionales.

De acuerdo con J. J. Heckman [1] el éxito académico está influido por una amplia gama de habilidades, debido a que, la calidad de la educación que reciben los niños en sus primeros años de vida son esenciales para su desarrollo cognitivo y socioemocional, ya que el cerebro humano es muy sensible a las primeras experiencias. Las habilidades socioemocionales como la persistencia y la autoestima son tan cruciales como las habilidades cognitivas, la comprensión lectora y las matemáticas, para predecir el éxito académico y profesional, además, menciona que el rendimiento académico en la escuela secundaria se relaciona con el éxito académico posterior, como la finalización de la educación superior y el desempeño en el mercado laboral [2].

Los autores C. Blair et.al [3] descubren que el rendimiento académico está vinculado al crecimiento cognitivo y socioemocional, como la capacidad de atención, la memoria de trabajo, las habilidades para resolver problemas y el control emocional. La capacidad de autorregulación es un factor predictivo más potente del rendimiento académico que el cociente intelectual (CI) o las capacidades cognitivas.

Existen diversas variables que pueden influir en el rendimiento académico de los estudiantes [4].

1. Factores individuales: Estos factores incluyen características individuales de los estudiantes, como su capacidad cognitiva, habilidades socio-emocionales, motivación y actitudes hacia el aprendizaje
2. Factores familiares: El apoyo y la implicación de los padres en la educación de sus hijos, así como el nivel socio-económico y educativo de la familia, también pueden influir en el rendimiento académico
3. Factores escolares: El ambiente escolar, la calidad de la enseñanza y la disponibilidad de recursos educativos también pueden influir en el rendimiento académico.
4. Factores culturales: La cultura y las normas sociales también pueden influir en el rendimiento académico.

B. *Machine Learning (ML)*

El ML es una rama de la inteligencia artificial que se enfoca en analizar, estudiar técnicas, desarrollar algoritmos y modelos para modelar el conocimiento a partir del aprendizaje. Esta tecnología ha sido posible gracias al desarrollo y a las capacidades tecnológicas que la humanidad ha adquirido con el tiempo. El ML tiene una amplia variedad de aplicaciones en diferentes campos, desde la medicina y la

industria hasta el comercio electrónico y la publicidad en línea. Se utiliza para el análisis y la predicción, automatización de procesos, reconocimiento de voz o imágenes, en el campo de la genética, seguridad informática, transporte autónomo, entre otros **2017una**.

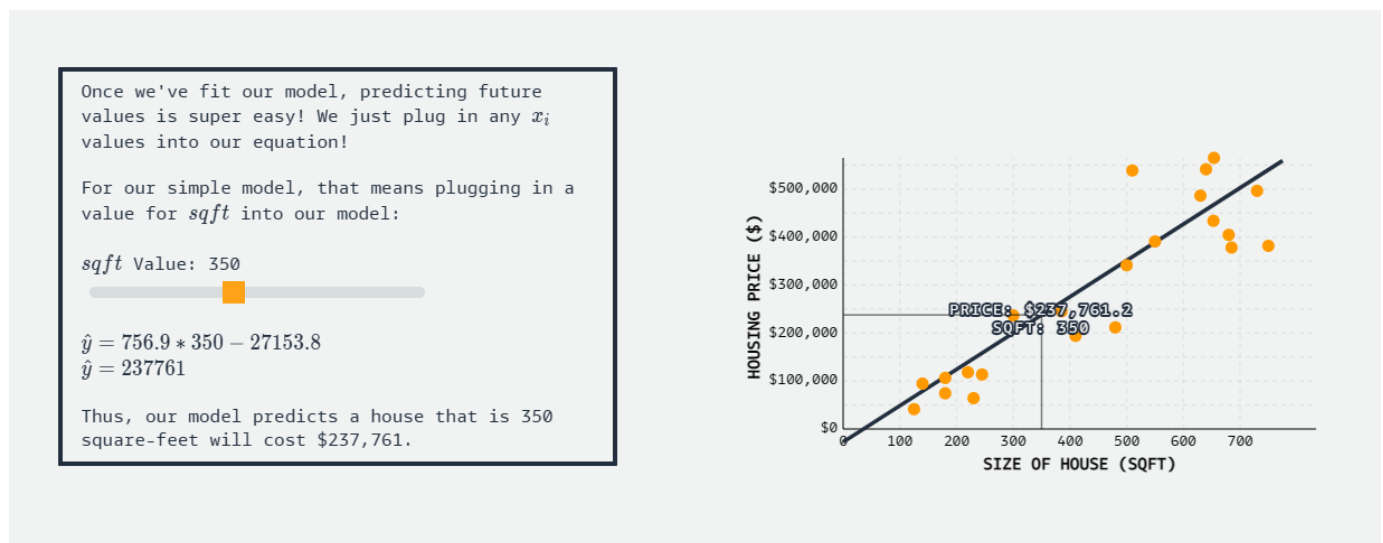
C. Tipos de algoritmos de ML

1) *Aprendizaje supervisado*: El aprendizaje supervisado es una técnica de ML en la que se entrena un modelo a partir de datos etiquetados. El objetivo del aprendizaje supervisado es encontrar una función que asigne de forma precisa y generalizable las entradas a las salidas correspondientes, es decir, encontrar una función que minimice el error entre las predicciones del modelo y las etiquetas reales de los datos de entrenamiento. El aprendizaje supervisado se utiliza en una amplia variedad de aplicaciones, desde la clasificación de imágenes y la detección de fraudes, hasta la predicción del precio de las acciones y la recomendación de productos. Los algoritmos de aprendizaje supervisado más comunes incluyen regresión lineal, los árboles de decisión, bosque aleatorio y las máquinas de vectores de soporte (SVM). La elección del algoritmo adecuado depende del tipo de problema que se esté abordando y de las características de los datos disponibles [5].

Algoritmos que se utilizan [6]:

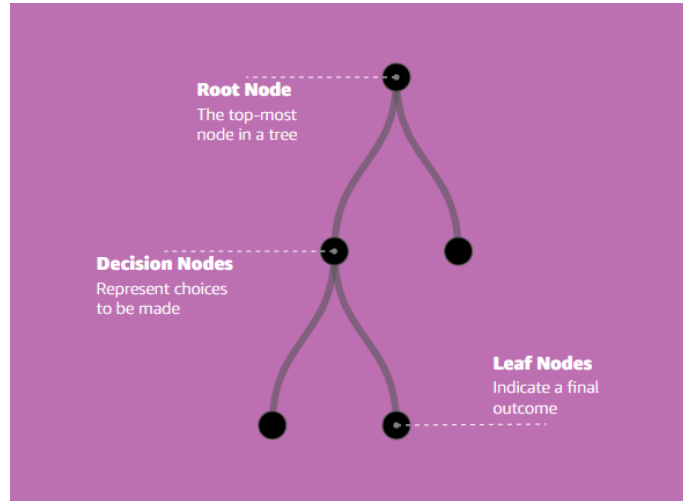
1. Regresión lineal: Es un algoritmo utilizado para predecir una variable continua a partir de una o más variables independientes. La regresión lineal busca encontrar la mejor línea recta que pueda ajustarse a los datos para realizar la predicción.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \epsilon \quad (1)$$



Donde y es la variable dependiente, x son las variables independientes, β son los coeficientes y ϵ es el error aleatorio.

2. Árboles de decisión: Son modelos de aprendizaje supervisado que se utilizan para clasificar o predecir a partir de múltiples variables. Los árboles de decisión construyen un árbol que representa las posibles decisiones a tomar para llegar a una predicción.



Donde root node es el nodo superior en un árbol, decision nodes representa las elecciones a realizar, las ramas indican las condiciones para llegar a un nodo determinado y leaf nodes indica un resultado final

3. Máquinas de vectores de soporte (SVM): Es un algoritmo utilizado para clasificar observaciones en dos o más categorías. SVM busca encontrar el hiperplano que mejor separa las categorías en un espacio de alta dimensión. Este hiperplano puede ser descrito por la ecuación:

$$w^T x + b = 0$$

Donde w es un vector normal al hiperplano y b es un sesgo (también conocido como término de sesgo o de intercepción).

La distancia entre un punto x y el hiperplano se puede calcular como:

$$d = \frac{|w^T x + b|}{||w||} \quad (2)$$

Donde $||w||$ es la norma euclidiana del vector w .

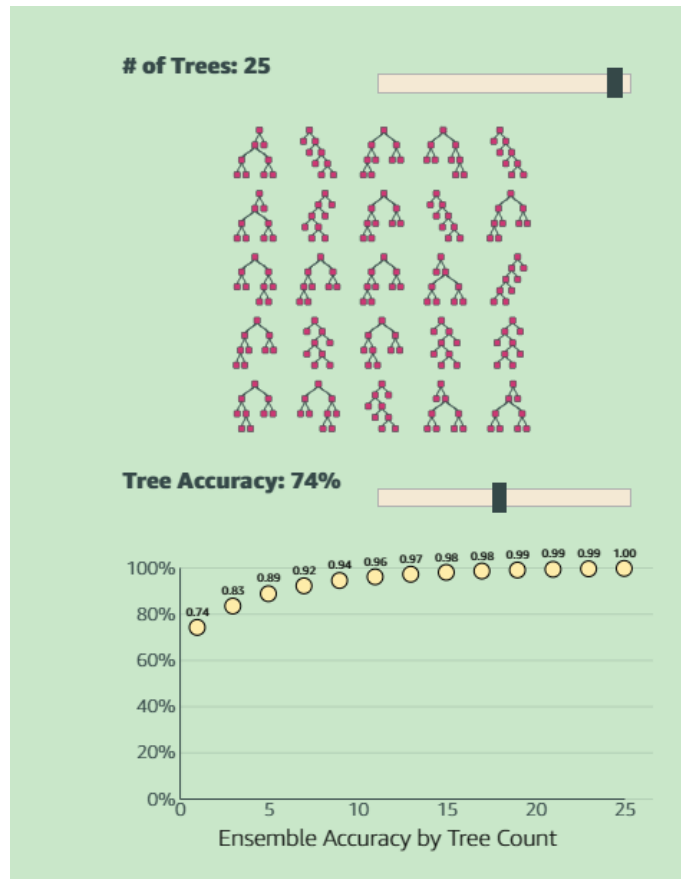
El objetivo de SVM es maximizar la distancia entre los dos márgenes de separación (uno para cada clase), lo que se puede expresar como:

$$\operatorname{argmax}_{w,b} \left(\frac{1}{||w||} \min_{i=1}^n [y_i(w^T x_i + b)] \right) \quad (3)$$

Donde argmax encuentra el argumento del máximo y el mínimo. Esto se puede reescribir como un

problema de optimización cuadrática con restricciones, que se puede resolver utilizando técnicas de optimización numérica.

4. Random Forest: Es un algoritmo utilizado en una amplia variedad de aplicaciones, como la detección de fraudes, la predicción de enfermedades, la clasificación de imágenes y el análisis de sentimiento. Es una técnica muy útil cuando se tienen datos con múltiples características y se desea obtener una predicción precisa y robusta. Además, combina varios árboles de decisión, de forma que se construye una “selva” o “bosque” de árboles de decisión.



La precisión del árbol de decisiones se refiere a la capacidad del modelo para hacer predicciones precisas, es decir, la precisión es la proporción de predicciones correctas que realiza el modelo en relación con el total de predicciones realizadas, y la cantidad de arboles nos permite tener más precisión.

2) *Aprendizaje no supervisado:* El aprendizaje no supervisado es una rama del ML que se enfoca en el análisis de datos sin la guía de variables de salida predefinidas o etiquetas. En este tipo de aprendizaje, el objetivo es encontrar patrones y estructuras en los datos, lo que puede reducir la complejidad de la información y proporcionar información valiosa.

El modelo se entrena utilizando técnicas como la agrupación, la reducción de la dimensionalidad y la asociación para identificar patrones y relaciones en los datos sin la necesidad de una guía explícita. Este

enfoque tiene varias aplicaciones, como la segmentación de clientes, el análisis de redes sociales y la detección de anomalías.

Algunos de los algoritmos de aprendizaje no supervisado más comunes son k-means clustering y análisis de componentes principales (PCA), los cuales se utilizan para encontrar patrones y relaciones ocultas en los datos [7].

3) *Aprendizaje por refuerzo*: El aprendizaje por refuerzo es una técnica de aprendizaje automático que permite a un agente tomar decisiones a través del proceso de "aprender haciendo". En este tipo de aprendizaje, el agente interactúa con el entorno y recibe información sobre el estado del mismo. El objetivo es que el agente elija un curso de acción que maximice una recompensa o minimice una penalización. La tarea del agente es desarrollar una política óptima que asigne el estado actual del entorno al curso de acción adecuado. El agente actualiza su política a medida que interactúa con el entorno, buscando maximizar la recompensa a largo plazo.

El aprendizaje por refuerzo tiene una amplia variedad de aplicaciones, como robótica, juegos y control de procesos. Por ejemplo, se puede utilizar para construir robots que aprendan a caminar o volar, o para optimizar procesos de producción y maximizar la eficiencia.

A diferencia del aprendizaje supervisado y no supervisado, el aprendizaje por refuerzo no utiliza un conjunto de datos etiquetados o no etiquetados para aprender. En su lugar, el agente aprende a través de la experiencia y la interacción con el entorno. Por lo tanto, el aprendizaje por refuerzo es especialmente útil en situaciones en las que el agente necesita aprender de forma autónoma y adaptarse a entornos dinámicos y cambiantes. En este proyecto, se explorarán algoritmos de predicción utilizando técnicas de aprendizaje por refuerzo para predecir el rendimiento académico de los estudiantes[8].

D. Estadística

1) *Error Cuadrático Medio (MSE)*: El Error Cuadrático Medio (MSE) es una medida estadística utilizada para cuantificar la discrepancia entre los valores observados y los valores predichos por un modelo. En el campo de la estadística y el aprendizaje automático, esta medida se utiliza ampliamente para evaluar la precisión de un modelo de predicción.

El cálculo del MSE se basa en la media de los errores cuadráticos entre los valores observados y los valores predichos:

$$MSE = \frac{1}{n} \sum_{i=0}^{n-1} (y_i - \hat{y}_i)^2$$

Donde n es el número de observaciones en el conjunto de datos, y_i es el valor real, \hat{y}_i es el valor estimado y \sum es la sumatoria.

El valor del MSE indica la calidad del ajuste del modelo. Un valor bajo del MSE indica una alta capacidad del modelo para hacer predicciones precisas, mientras que un valor alto del MSE indica una baja capacidad del modelo para hacer predicciones precisas [9].

2) *Error Absoluto Medio (MAE)*: El Error Absoluto Medio (MAE) es una medida utilizada habitualmente en estadística para evaluar la precisión de los modelos de regresión ML. El MAE es una medida importante para evaluar la precisión de un modelo de regresión. También es especialmente útil en situaciones en las que se desea minimizar los efectos de los valores atípicos, ya que mide la diferencia absoluta entre los valores reales y los predichos. El calculo del MAE se basa en el promedio de la diferencia absoluta entre los valores reales y los valores estimados:

$$MAE = \frac{1}{n} \sum_{i=0}^{n-1} |y_i - \hat{y}_i|$$

Donde n es el número de observaciones en el conjunto de datos, y_i es el valor real, \hat{y}_i es el valor estimado y \sum es la sumatoria.

El valor del MAE representa el error absoluto medio de las predicciones del modelo de regresión, es decir, el valor medio de la diferencia absoluta entre las predicciones del modelo y los valores reales de la variable de interés. Un MAE más bajo indica que las predicciones del modelo son más precisas porque la diferencia entre los valores predichos y los reales es menor por término medio [10].

3) *Raiz del Error Cuadrático Medio (RMSE)*: La Raiz del Error Cuadrático Medio (RMSE) es una medida utilizada habitualmente para evaluar la precisión de los modelos de regresión ML. Esta medida se calcula como la raíz cuadrada del promedio de las diferencias cuadráticas entre los valores reales y los valores estimados. RMSE se utiliza a menudo cuando se desea dar más peso a los grandes errores en un pronóstico, como cuando se predicen valores atípicos. También puede resultar útil para comparar la precisión de distintos modelos de regresión.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Donde n es el número de observaciones en el conjunto de datos, y_i es el valor real, \hat{y}_i es el valor estimado y \sum es la sumatoria.

El valor de RMSE indica la diferencia media entre las predicciones de un modelo y el valor real de la variable de interés. Como medida de error, cuanto menor sea el valor de RMSE, más acertadamente predice el modelo la variable de interés. Y como se ha mencionado anteriormente, el RMSE es especialmente útil cuando se desea penalizar errores grandes en las predicciones [10].

VII. METODOLOGÍA

A. Descripción del dataset.

Para la realización de este proyecto se evaluó en primera instancia la implementación de la base de datos de la universidad para que se pudiera aplicar en el contexto educativo en el cual se está cursando, pero debido a los permisos requeridos para la obtención de estos datos y posibles conflictos con la protección de los datos de los estudiantes, se optó por utilizar un dataset publico relacionado al rendimiento académico estudiantil. El conjunto de datos "Student Performance"(Rendimiento de los estudiantes) es un conjunto de datos recopilado a partir de registros escolares y encuestas realizadas en dos escuelas secundarias de Portugal. Proporciona información detallada sobre el rendimiento académico de los estudiantes en el curso de matemáticas en dos escuelas secundarias de Portugal.

Este conjunto de datos se encuentra disponible en el Repositorio de Aprendizaje Automático de la Universidad de California Irvine (UCI Machine Learning Repository) [23], una fuente confiable y ampliamente utilizada para conjuntos de datos de aprendizaje automático y análisis de datos. Además, está disponible de forma gratuita para fines de investigación y se proporciona con el objetivo de fomentar el estudio del rendimiento estudiantil y los factores que pueden influir en él.

El dataset que se seleccionó incluye un total de 33 variables con un total de 395 datos iniciales, incluyendo información sobre las características demográficas, los hábitos de estudio de los estudiantes, antecedentes familiares y otros factores que pueden influir en su rendimiento académico. Adicionalmente, estos datos pueden ser utilizados para diversos propósitos, como analizar los factores que contribuyen al éxito académico, detectar conductas que llevan a la deserción académica o identificar a los estudiantes en riesgo que puedan necesitar apoyo adicional.

1) Atributos Generales.:

- 'School': Nombre de la escuela a la que pertenece el estudiante ("GP." "MS").
- 'Sex': Género del estudiante ("F"para femenino o "M"para masculino).
- 'Age': Edad del estudiante (numérico).
- 'Address': Tipo de dirección del estudiante (Ü"para urbana o R"para rural).
- 'Famsize': Tamaño de la familia del estudiante ("LE3"para menos de 3 o "GT3"para igual o más de 3).
- 'Pstatus': Estado de convivencia de los padres ("T"para conviven o .^"para separados).

2) Atributos Educativos.:

- 'Medu': Nivel de educación de la madre (0-4).
- 'Fedu': Nivel de educación del padre (0-4).
- 'Traveltime': Tiempo de viaje desde el hogar a la escuela (1-4).
- 'Studytime': Tiempo de estudio semanal (1-4).

- 'Failures': Número de clases anteriores reprobadas (0-3).
- 'Schoolsup': Apoyo educativo adicional de la escuela (ses.º "no").
- 'Famsup': Apoyo educativo adicional de la familia (ses.º "no").
- 'Paid': Clases particulares pagadas (ses.º "no").
- 'Activities': Actividades extracurriculares (ses.º "no").
- 'Internet': Acceso a internet en el hogar (ses.º "no").

3) *Atributos Sociales.:*

- 'Famrel': Calidad de las relaciones familiares (1-5).
- 'Freetime': Tiempo libre fuera de la escuela (1-5).
- 'Goout': Salidas con amigos (1-5).
- 'Dalc': Consumo de alcohol en días laborales (1-5).
- 'Walc': Consumo de alcohol en fines de semana (1-5).
- 'Healt': Estado de salud actual (1-5).

4) *Atributos relacionados con el rendimiento escolar.:*

- 'Absences': Número de ausencias escolares (0-93).
- 'G1': Calificación en el primer período (0-20).
- 'G2': Calificación en el segundo período (0-20).
- 'G3': Calificación final en matemáticas (variable objetivo) (0-20).

B. *Metodología*

Para este estudio se optó por la metodología de investigación descriptiva teniendo en cuenta que se utilizan datos específicos para el diseño de los modelos concebidos a partir de la información recolectada, adicional utilizar el modelo de cascada como metodología de trabajo, esta se divide en 5 actividades y deben ser ejecutadas en orden secuencial ya que cada una depende de la otra.

1) *Determinar los modelos predictivos:* En esta fase se debe investigar y recolectar todos los algoritmos predictivos existentes. asimismo, recolectar la información necesaria acerca de las técnicas de minería de datos para un posterior análisis a cada uno de ellos.

2) *Evaluar los modelos predictivos:* Esta actividad va muy relacionada a la anterior dado que aquí se evalúa a detalle las fuentes de donde se tomó dicha información asimismo la revisión de los antecedentes de los algoritmos predictivos ya preseleccionados.

3) *Establecer el mejor modelo:* Una vez seleccionado los algoritmos de machine learning se debe determinar cuál de todos es el que considera más adecuado para el tratamiento de los datos. Adicional a esto se deben llevar acabo las siguientes actividades:

- Determinar los algoritmos de machine learning que se consideren adecuados para el tratamiento de los datos.
- Obtención de los datos iniciales.
- Descripción de los datos.
- Verificar la calidad de los datos.

4) *Ejecutar algoritmos de aprendizaje:* Aquí se debe poner en práctica y ejecutar los algoritmos seleccionados, experimentando con diferentes configuraciones e identificar los mejores modelos mediante la comparación de los resultados. Adicional a esto se deben llevar a cabo las siguientes actividades:

- Realizar el cargue de los datos a la plataforma de minería de datos.
- Realizar el preprocesamiento de los datos.
- Integrar los datos a partir de la combinación de las fuentes de información
- Selección de los datos aplicando criterios de inclusión y exclusión
- Identificación de los mejores modelos mediante la comparación de los resultados.
- Limpieza de datos.

5) *Análisis de datos obtenidos:* Analizar y redactar las conclusiones obtenidas. Aquí debemos tener presente los resultados de los algoritmos ejecutados y los indicadores de precisión según los modelos. En esta fase se llevan a cabo las siguientes actividades:

- Análisis de los resultados de los algoritmos ejecutados.
- Elaboración del informe final del proyecto
- Elaboración de conclusiones.

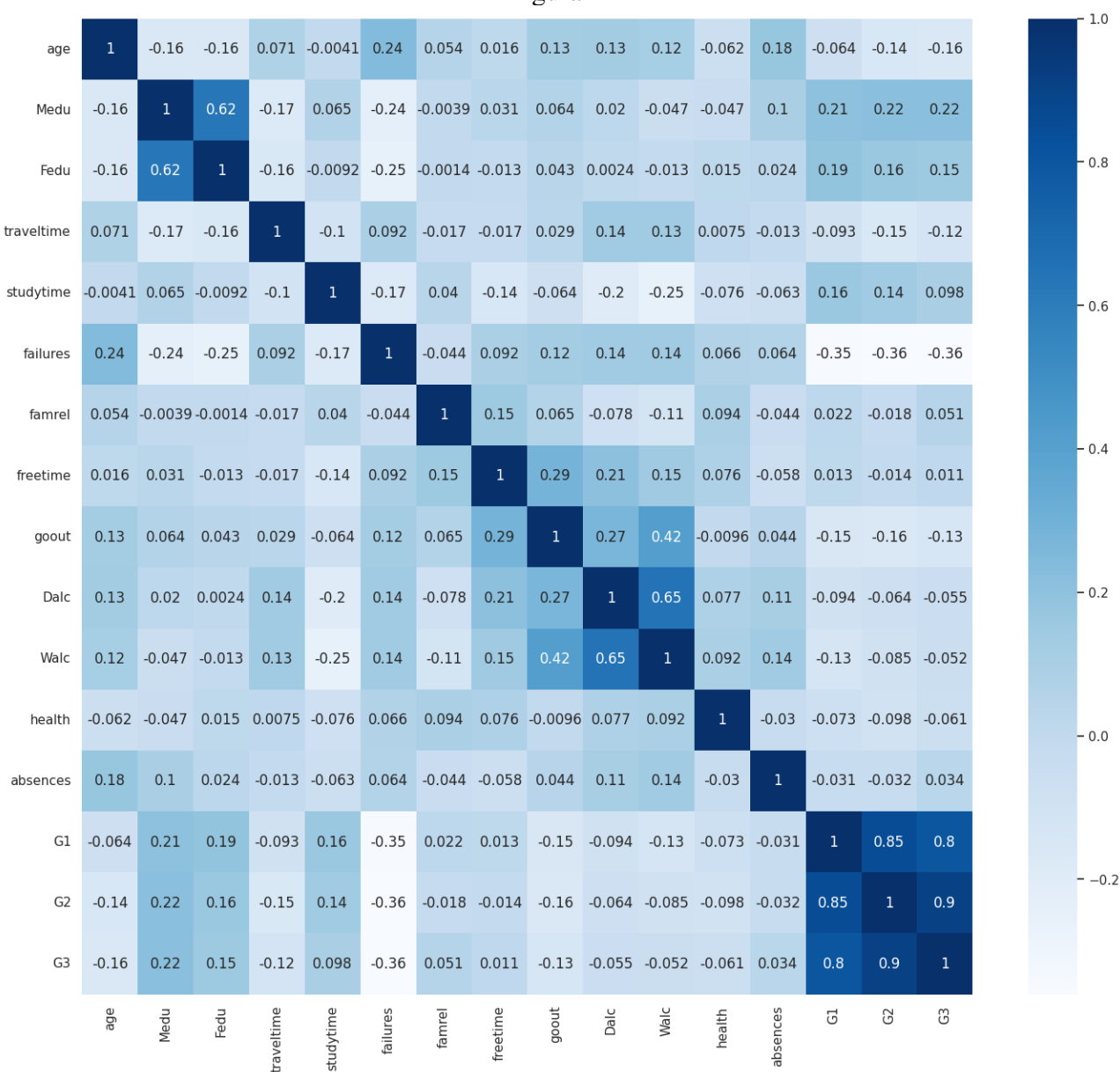
VIII. ANÁLISIS EXPLORATORIO DE DATOS

Con este tipo de análisis exploratorio, se puede obtener información valiosa sobre los factores que influyen en el rendimiento académico de los estudiantes y ayudar a las instituciones educativas a tomar decisiones informadas para mejorar su desempeño y éxito académico.

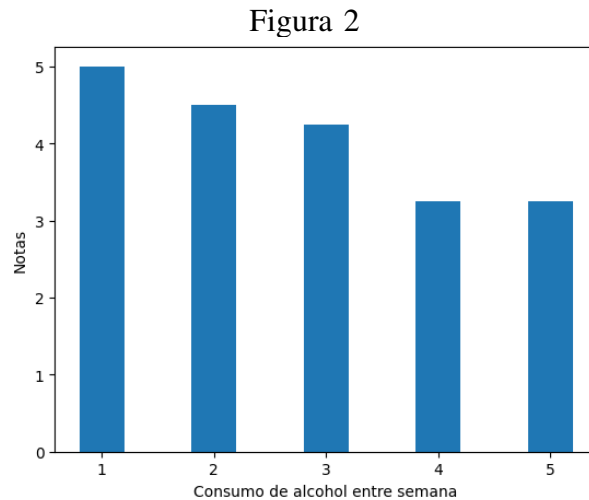
1. Comprensión de los datos.

1.1. A partir de la matriz de correlación se pueden explorar patrones y relaciones en un conjunto de datos multivariable. Puede ayudar a identificar variables que están altamente correlacionadas entre sí, lo que puede indicar una dependencia o influencia mutua. Además puede ser utilizada para seleccionar características relevantes en el análisis de datos o para detectar problemas como la multicolinealidad en modelos estadísticos.

Figura 1

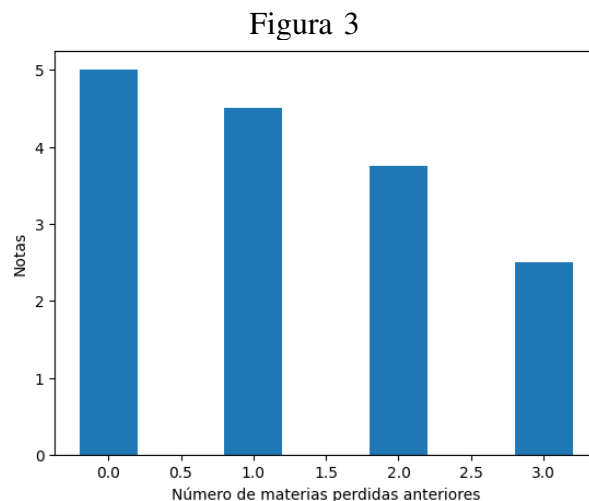


- Basándonos en el análisis de la Figura 1. con respecto a la correlación lineal existente entre las variables de todo el conjunto de datos, se puede evidenciar como algunas se encuentran estrechamente vinculadas, por lo que aquellas que tienen el cuadro relleno y completamente azul indican una correlación de 1 lo que permite identificar qué variables están altamente conectadas, Una ilustración notable de lo previamente mencionado es comparar la nota final(G3) con el consumo de alcohol entre semana(Dalc)



En la Figura 2. se observa que a medida que el consumo de alcohol por parte del estudiante aumenta durante la semana, la probabilidad de obtener una calificación baja también se incrementa.

Continuando con este mismo patrón se analiza la variable nota final(G3) con número de materias perdidas en los semestres pasados(failures)



Mediante este tipo de análisis, se puede determinar qué variables son fundamentales para que un estudiante logre una calificación satisfactoria o no.

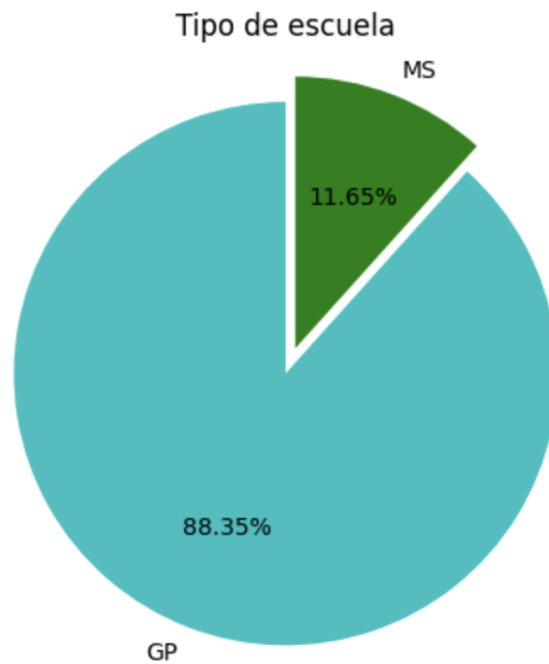
1.2. Variables categóricas

1.2.1. First Sub-Subtopic

1.2.2. Second Sub-Subtopic

■ school:

Tipo de escuela a la que asiste el estudiante (binaria: 'GP' - Gabriel Pereira o 'MS' - Mousinho da Silveira) Esta variable es categórica nominal binaria.

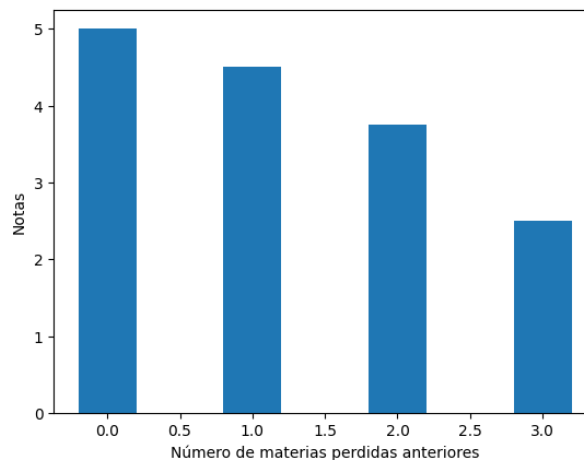


■ sex:

Género del estudiante (binario: 'F' - femenino o 'M' - masculino).

■ age:

Edad del estudiante (numérica: de 15 a 22) Esta variable es numérica discreta.



■ school:

Ubicación de la casa del estudiante (binaria: 'U' - urbana o 'R' - rural) Esta variable es categórica

nominal binaria.

- famsize:

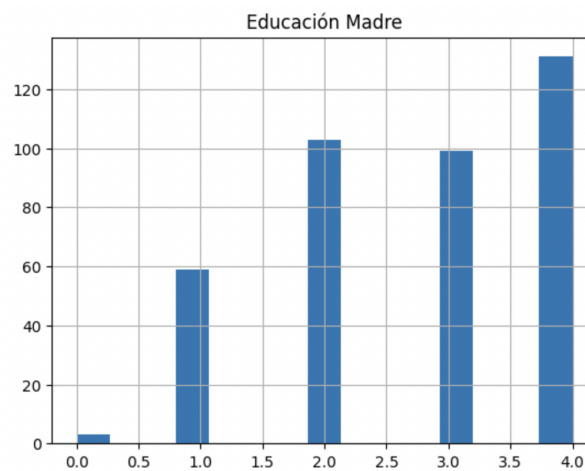
Tamaño de la familia del estudiante (binaria: 'LE3' - ≤ 3 o 'GT3' - > 3). La mayoría de los estudiantes provienen de familias más grandes que 3 miembros

- Pstatus:

Estado de convivencia de los padres del estudiante (binaria: 'T' - viven juntos o 'A' - viven separados). En el conjunto de datos se puede observar que la mayoría de los estudiantes provienen de familias donde los padres viven juntos.

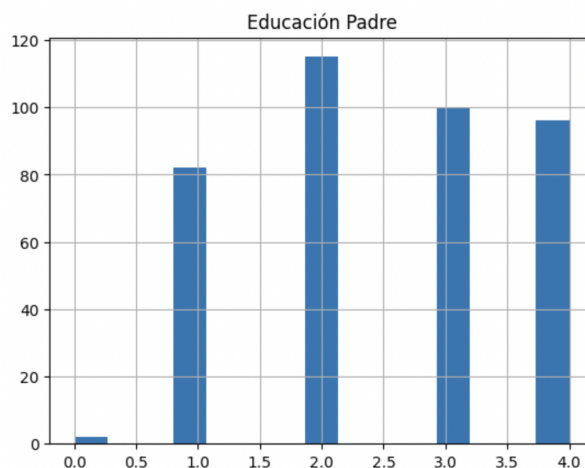
- Medu:

Nivel de educación de la madre (numérica: 0 - sin educación, 1 - educación primaria (4° grado), 2 - educación de 5° a 9° grado, 3 - educación secundaria, 4 - educación superior)



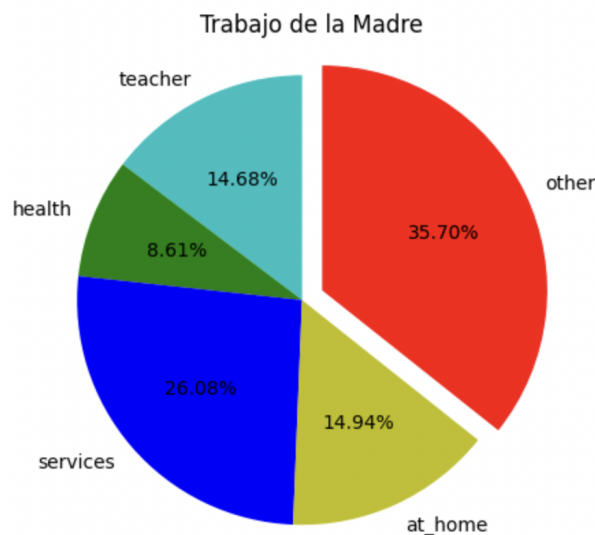
- Fedu:

Nivel de educación del padre (numérica: 0 - sin educación, 1 - educación primaria (4° grado), 2 - educación de 5° a 9° grado, 3 - educación secundaria, 4 - educación superior)



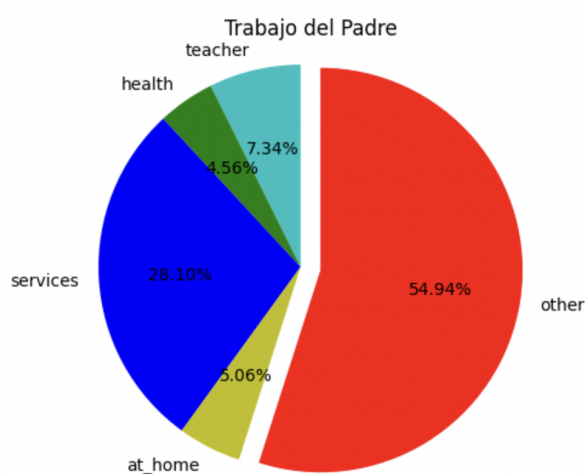
- Mjob:

Trabajo de la madre (nominal: 'teacher', 'health'(relacionado con el cuidado), 'services' trabajos civiles (Administrativa o policial), 'at_home' or 'other')



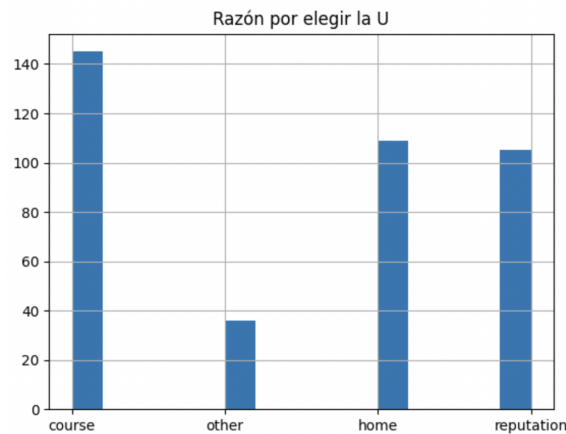
■ Fjob:

Trabajo del padre (nominal: 'teacher', 'health'(relacionado con el cuidado), 'services' trabajos civiles (Administrativa o policial), 'at_home' or 'other')



■ reason:

Razón para elegir esta U (nominal: cerca de 'home', reputación de la escuela 'reputation', preferencia de curso 'course' o 'other') Esta variable es categórica nominal.



- guardian:

Tutor legal del estudiante (nominal: 'mother', 'father' o 'other'). De los cuales 273 tienen a su madre como tutor legal, 90 tienen a su padre y 32 tienen a otra persona como tutor legal.

Se puede evidenciar que la mayoría de los estudiantes tienen a su madre como tutor legal.

- traveltime:

Tiempo de viaje de casa a la escuela (numérica: 1: ≤ 15 min, 2: 15 a 30 min, 3: 30 min a 1 hora, 4: ≥ 1 hora) Esta variable es numérica discreta.

La mayoría de los estudiantes tienen un tiempo de viaje de menos de 30 minutos.

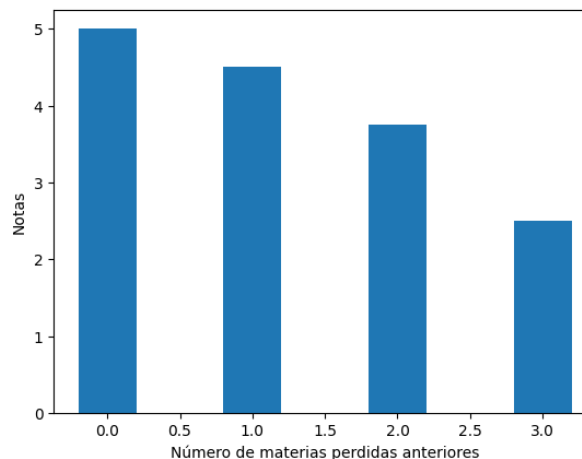
- studytime:

Tiempo semanal de estudio (numérica: 1: ≤ 2 horas, 2: 2 a 5 horas, 3: 5 a 10 horas, 4: ≥ 10 horas)

La mayoría de los estudiantes dedican entre 2 y 5 horas a estudiar por semana.

- failures:

Número de fallos de clases anteriores (numérica: n si $1 \leq n \leq 3$, de lo contrario 4).



- schoolsup:

Soporte educativo extra de la escuela (binaria: sí o no). La mayoría de los estudiantes no reciben apoyo educativo extra de la escuela.

- famsup:

Soporte educativo extra de la familia (binaria: sí o no).

- paid:

Clases particulares pagadas (binaria: sí o no).

- activities:

Actividades extracurriculares (binaria: sí o no). Aquí se puede evidenciar que la mayoría de los estudiantes no participan en actividades extracurriculares.

- nursery:

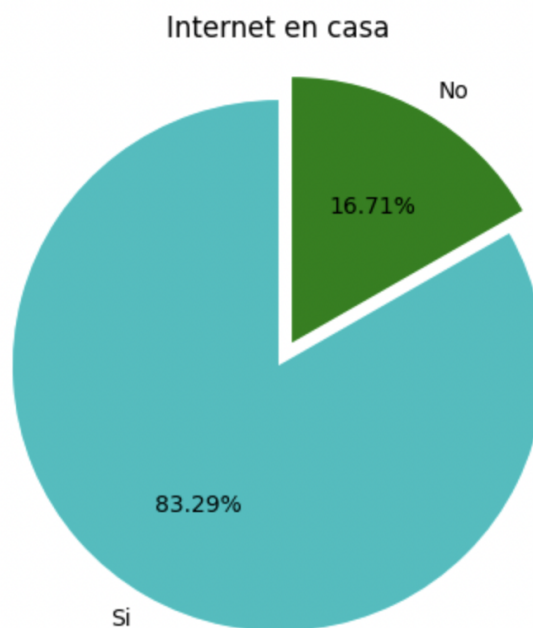
Asistió a guardería infantil (binaria: sí o no). La mayoría de los estudiantes asistieron a guardería infantil.

- higher:

Quiere seguir educación superior (binaria: sí o no). La mayoría de los estudiantes quieren seguir educación superior.

- internet:

Acceso a Internet en casa (binaria: sí o no).



- romantic:

Relación romántica (binaria: sí o no). La mayoría de los estudiantes no tienen una relación romántica.

- famrel:

Calidad de las relaciones familiares (numérica: de 1 - muy malo a 5 - excelente).

IX. CONCLUSIONES

Son las interpretaciones finales que recopilan los datos de la investigación, describe lo que se obtuvo, qué se logró y cuáles son los resultados. Guardan relación directa con lo que se mencionó en el planteamiento del problema. Pueden confirmar las hipótesis.

X. RECOMENDACIONES

No obligatorio Las recomendaciones son las futuras y posibles líneas de investigación que llevarán a resolver problemas relacionados con la presente investigación.

REFERENCIAS

- [1] J. J. Heckman, «Schools,Skills,Synapses,» 2008.
- [2] P. Carneiro, C. Crawford y A. Goodman, «“Which Skills Matter?”»,» 2006.
- [3] C. Blair y R. P. Razza, «Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten,» 2007.
- [4] É. C. Vargas y G. R. Villalobos, «Factors affecting the academic performance of freshmen in the ITCR general mathematics course,» 2021.
- [5] J. G. Cangrejo, *Algoritmos de Aprendizaje Supervisado en la Clasificación de Exoplanetas en Python*. 2021.
- [6] D. Patiño-Pérez, F. Iñiguez-Muñoz, M. Nivelá-Cornejo et al., *Machine Learning models based in Supervised Learning for the Detection of Diabetes Mellitus in the City of Guayaquil*. 2023.
- [7] J. L. Melgarejo, A. T. Trigos, M. C. Cara y D. Miranda, «Identification of risk zones for road safety through unsupervised learning algorithms — Identificación de zonas de riesgo para la Seguridad Vial mediante algoritmos de aprendizaje no supervisado,» 2018.
- [8] J. E. S. García y M. Santos, «Neural networks and reinforcement learning in wind turbine control,» 2021.
- [9] T. O. Hodson, T. M. Over y S. S. Foks, «Mean Squared Error, Deconstructed,» 2021.
- [10] scikit-learn, «scikit-learn: machine learning in Python — scikit-learn 1,”2. 2 documentation”. scikit-learn: machine learning in Python — scikit-learn 0. 16,»

APÉNDICE A

EJEMPLO DE ANEXO: DATOS

Un ejemplo de información que debe ir en la base en su tesis para poder reproducirla, pero que no hace parte de tu trabajo son los apéndices.

APÉNDICE B

OTRO APÉNDICE

Imágenes de modelos finales, datos técnicos de equipos, Definiciones grandes de modelos matemáticos son información adicional que pueda ir en los apéndice.