

# PRODUCTIVIDAD EDUCATIVA Y ALGORITMOS DE APRENDIZAJE AUTOMÁTICO EN ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA, UNIVERSIDAD NACIONAL DE MOQUEGUA- 2023

Wildon Rojas Paucar

Alberto Octavio Carranza López

Alberto Limache Flores

Elena Miriam Chávez Garcés

Elvis Alberto Pareja Granda



Año 2025

**Productividad educativa y algoritmos de aprendizaje automático en  
escuela de ingeniería de sistemas e informática, Universidad Nacional  
de Moquegua- 2023**

Wildon Rojas Paucar

ORCID: <https://orcid.org/0000-0001-6590-3225>

wrojas@unam.edu.pe

Universidad Nacional de Moquegua

Perú

Alberto Octavio Carranza López

ORCID: <https://orcid.org/0000-0002-5147-2134>

acarranzal@unmsm.edu.pe

Universidad Nacional Mayor de San Marcos

Perú

Alberto Limache Flores

ORCID: <https://orcid.org/0000-0003-1302-9569>

alimachef@unam.edu.pe

Universidad Nacional de Moquegua

Perú

Elena Miriam Chávez Garcés

ORCID: <https://orcid.org/0000-0002-0384-8758>

echavezg@unjbg.edu.pe

Universidad Nacional Jorge Basadre Grohmann

Perú

Elvis Alberto Pareja Granda

eparejag@unam.edu.pe

ORCID: <https://orcid.org/0000-0002-1266-7115>

Universidad Nacional de Moquegua

Perú

Productividad educativa y algoritmos de aprendizaje automático en escuela de ingeniería de sistemas e informática, Universidad Nacional de Moquegua- 2023

© 2025 Autores. Reservados todos los derechos.

Autores: Wildon Rojas Paucar, Alberto Octavio Carranza López, Alberto Limache Flores, Elena Miriam Chávez Garcés, Elvis Alberto Pareja Granda

Edición: Omniscens

Diseño de cubierta: Omniscens

Diseño interior: Omniscens

Primera edición: 2025

ISBN: 978-970-96928-5-3

DOI: <https://doi.org/10.7112/gm4g7278>

Sello editorial: Omniscens Publishing (97897096928)

Categoría: Computación

Tipo de Contenido: Libros Universitarios

Lugar de publicación: Mérida, Yucatán, México

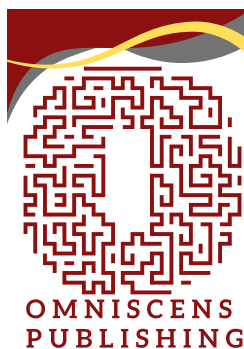
Esta obra está bajo una licencia internacional Creative Commons Atribución 4.0.

<https://creativecommons.org/licenses/by/4.0/>



**Omniscens Publishing**

[www.publishing.omniscens.com](http://www.publishing.omniscens.com)



## **DEDICATORIA**

A nuestras familias, por su amor y apoyo incondicional durante el proceso de investigación con atributos de paciencia, dedicación e inspiración en función del tiempo.



## **AGRADECIMIENTO**

A quienes apoyaron e hicieron posible realizar el presente trabajo de investigación, cuya contribución ha sido fundamental en su desarrollo y para el avance del conocimiento en general.

## **RECONOCIMIENTO**

Un reconocimiento especial a la Universidad Alas Peruanas, asesores, docentes, directivos de la institución, quienes han contribuido a la realización del presente trabajo de investigación. Su apoyo y colaboración han sido esenciales para el éxito respectivo.

# ÍNDICE

DEDICATORIA	3
AGRADECIMIENTO	4
RECONOCIMIENTO	5
RESUMEN	7
INTRODUCCIÓN	8
CAPÍTULO 1 Planteamiento del problema	10
CAPÍTULO 2: Marco filosófico	21
CAPÍTULO 3: Marco teórico conceptual.	34
CAPÍTULO 4: Hipótesis, variables y metodología de la investigación	71
CAPÍTULO 5: Resultados	82
ANEXOS	153

## RESUMEN

La presente investigación tiene como propósito determinar el nivel de relación entre la productividad educativa y los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023, con un tipo de investigación básica, de nivel descriptivo correlacional, transversal; siendo de diseño no experimental, método hipotético deductivo con enfoque cuantitativo.

La muestra estuvo conformada por 90 estudiantes de un total de 116, se realizó un muestreo aleatorio simple, como técnica se empleó la encuesta para la variable productividad educativa y la observación para la variable algoritmos de aprendizaje automático, con instrumento de cuestionario y lista de cotejo respectivamente con escala de Likert para las variables productividad educativa y algoritmos de aprendizaje automático, los mismos que fueron validados previamente demostrando la eficacia y confiabilidad mediante la técnica de opinión de expertos y alfa de Cronbach cuyo resultado fue 0.777 y 0.974 con significado de confiabilidad muy bueno y excelente.

Los resultados que se obtuvieron evidenciaron, que la productividad educativa y los algoritmos de aprendizaje automático tiene correlación positiva muy alta, debido a que el coeficiente de correlación fue 0,940 en la Escuela Profesional de Ingeniería de Sistemas e Informática de la Universidad Nacional de Moquegua, 2023, con un nivel de significancia es  $0,000 < 0,05$ , de acuerdo a la encuesta se determinó el nivel de relación que existe entre las variables productividad educativa y los algoritmos de aprendizaje automático fue 113.3% (12) Insuficiente, 71.1% (64) Aprobado y 15.6% (14) Bueno.

**Palabras clave:** Productividad académica, algoritmos de aprendizaje automático, resultados de aprendizaje, confiabilidad.

## INTRODUCCIÓN

El aprendizaje automático (ML) está de moda, la era de la digitalización genera la producción masiva de datos y esto se debe a las aplicaciones automatizadas, el incremento de la potencia de cálculo y el desarrollo de nuevos algoritmos; hoy en día el aprendizaje automático se aplican en diferentes campos, los diferentes sectores intentan beneficiarse con conocimientos inteligentes, los dispositivos móviles ya cuentan con dicha tecnología, y es muy posible se encuentre en la mayoría de nuestras manos, sus aplicaciones están diversificadas en predicciones, reconocimiento de imágenes, reconocimiento de voz, utilizadas en finanzas, comercio, medicina, educación, etc.

En el campo educativo universitario, es un hecho que los procesos académicos están digitalizados, por tanto, utilizar diferentes técnicas de ML que proporcionen estimaciones sobre fenómenos, que realizan recomendaciones es significativo para tomar decisiones; cada año el ingreso de nuevos estudiantes a las universidades, es un logro que genera grandes expectativas en la sociedad, para ello la productividad académica basado en el rendimiento académico del estudiante es el resultado más importante dentro de las experiencias de la vida universitaria, su importancia se basa en el grado alto o bajo desempeño que afecta al estudiante y por ende a la universidad (Garbanzo 2014 ,como se citó en Fernández-Mellizo y Constante-Amores, 2020); asimismo, se establece que el rendimiento académico refleja el proceso de aprendizaje ligado a la capacidad y esfuerzo del estudiante, a medidas de calidad y juicios de valoración que posteriormente permitirá alcanzar expectativas sociales ( García y Palacios, 1991, como se citó en Alban y Calero J. L., 2017). Conocer la productividad del rendimiento académico es vital para toda institución educativa, significa conocer el nivel cognitivo del estudiante que adquirió durante el proceso enseñanza aprendizaje conforme a los lineamientos del plan de estudios de las escuelas profesionales.

El escenario es significativo para las instituciones educativas de nivel superior debido a la gran cantidad de datos académicos existentes, por ello predecir el desempeño académico es complicado, por tal motivo el presente estudio parte de los diferentes trabajos de investigación descritas en los antecedentes del problema en donde se analiza los diferentes indicadores que influyen en el rendimiento académico, acordes a las realidades académicas de los estudiantes permitiendo conocer el estado académico del estudiante; para ello el presente trabajo de investigación está orientado a predecir la productividad académica utilizando algoritmos de clasificación de aprendizaje automático: Árboles de decisión, Naive Bayes y Random Forest pertenecientes a la corriente de la Inteligencia Artificial, cuyos



algoritmos de aprendizaje automático permitirán aprender y describir la productividad del rendimiento académico considerando escenarios e indicadores de actividades de estudio, adquisición de información, aspectos comunitarios y aspectos de desempeño académico; así de esta manera podamos adquirir nuevos conocimientos de clasificación, aprendizaje y eficiencia comparativa de los algoritmos mencionados acordes a nuestra realidad, los mismo que establecerán información colaborativa para la toma de decisiones correspondientes en la gestión académica de los estudiantes de la Escuela Profesional de Ingeniería de Sistemas e Informática (EPISI) de la Universidad Nacional de Moquegua (UNAM), cuyos datos académicos conciernen al periodo académico 2023.

# CAPÍTULO 1

## Planteamiento del problema

### Problem statement

DOI: <https://doi.org/10.71112/kv7zdx52>

#### Resumen

El capítulo plantea el problema de la baja integración del aprendizaje automático en la gestión académica de la EPISI de la Universidad Nacional de Moquegua, pese a la digitalización de sus procesos. Se identifica que no se aplican adecuadamente algoritmos como árboles de decisión, Naive Bayes y Random Forest para analizar la productividad educativa basada en el rendimiento académico. La investigación busca determinar la relación entre productividad educativa y aprendizaje automático, justificando su relevancia teórica, práctica, metodológica y social para mejorar la calidad educativa universitaria

**Palabras clave:** Productividad educativa, aprendizaje automático, rendimiento académico, algoritmos, educación superior.

#### Abstract

This chapter addresses the problem of the low integration of machine learning into the academic management of the EPISI at the National University of Moquegua, despite the digitalization of its processes. It identifies the inadequate use of algorithms such as decision trees, Naive Bayes, and Random Forest to analyze educational productivity based on academic performance. The research aims to determine the relationship between educational productivity and machine learning, justifying its theoretical, practical, methodological, and social relevance to improve the quality of university education.

**Keywords:** Educational productivity, machine learning, academic performance, algorithms, higher education.

#### Descripción de la realidad problemática

El nuevo escenario digital denominado cuarta revolución industrial ha transformado las sociedades, no cabe duda que ha alterado la forma de vivir y laboralmente trabajamos de una manera distinta, las nuevas tecnologías, como la computación en la nube, internet de las cosas, inteligencia artificial, etc., se ha integrado a nuestras actividades de manera directa o indirecta,

esto significa que estamos digitalizados (McGinnis, 2020); por ello la digitalización académica no es ajena en UNAM, los procesos están automatizados, esto significa que se encuentran estructurados según el modelo de cada escuela profesional, para nuestro caso *EPISI*, en donde *los procesos automatizados generan información distribuida sin tomar decisiones significativas concernientes a la productividad educativa basados en el rendimiento académico*, es decir a partir de los datos tenemos conocimiento de la realidad académica de los actores involucrados pero no existe una herramienta de aprendizaje automático que analice de manera inteligente el comportamiento de los datos a fin de modelar y evaluar la productividad educativa de los estudiantes mientras van reconociendo debilidades y se van mejorando los resultados educativos desde diferentes enfoques como los algoritmos de aprendizaje automático Árboles de decisión, Naive Bayes y Random Forest.

Por ello, la revisión realizada *destaca* que la generación masiva de datos ha estimulado a la inteligencia artificial a través de los diferentes algoritmos funcionales a aprender de nuestras actividades, asimismo se ha adaptado a nuestras actividades cognitivas; considerando a Sarker (2021) en donde menciona que para analizar inteligentemente los datos y desarrollar aplicaciones inteligentes, es importante aplicar el conocimiento de la inteligencia artificial (IA) particularmente el aprendizaje automático (ML), asimismo según Edgard y Manz (2017), menciona que el aprendizaje automático es el campo de la informática que se encarga de enseñar a las computadoras cómo aprender de los datos, dicho aprendizaje depende del algoritmo que es un conjunto de reglas, en donde indica lo que debe realizar una computadora; Ray (2019) menciona algunos algoritmos de aprendizaje automático más utilizados con mayor frecuencia y populares: Algoritmo de descenso de gradientes, Algoritmo de regresión lineal, Análisis de regresión multivariado, Regresión logística, Árbol de decisión, Máquinas de vectores soporte, Aprendizaje bayesiano, Naïve bayes,c, K means algoritmo de clustering, Algoritmo de propagación de espalda, etc., *dichos algoritmos de aprendizaje automático que en los diferentes ámbitos empresariales han sido muy importantes*, asimismo dichas aplicaciones han generado aportes significativos en la toma de decisiones considerando el nivel de confianza que brindan las diferentes técnicas.

Mediante Ley 30220, SUNEDU en su art. 1, destaca la “promoción del mejoramiento continuo de la calidad educativa de las instituciones universitarias como entes fundamentales del desarrollo nacional, de la investigación y de la cultura”, el tema de calidad educativa universitaria es un punto crucial que ha enfocado SUNEDU en las universidades del país, por tanto, la productividad educativa a través del rendimiento académico es una variable importante de eficiencia y calidad en la educación superior; es aquí en donde los datos juegan

un rol importante a fin de generar información relevante respecto a lo mencionado, según (Hawkins y Bailey, 2020) destaca la importancia del uso y el valor de los datos de los entes que componen la universidad a fin de generar información destacable para diferentes propósitos y niveles de la administración académica; asimismo, a fin de medir la productividad educativa, se debe definir los indicadores necesarios con la finalidad de obtener el producto esperado (Hanushek y Ettema, 2017, como se citó en Triwiyanto T. et al., 2020); por tanto la calidad educativa basado en los datos y cómo van generando valor desde los diferentes niveles establecerán información crucial e importante para el conocimiento de la productividad educativa, cabe recordar que la productividad educativa de los estudiantes es importante en todo sistema educativo, los indicadores influyentes deben cumplir aspectos influyentes a fin de obtener información relevante (Brew et al., 2021); según, Medina N. et al. (2018), recalcan que la medición del rendimiento académico y su predicción considerando las variables que lo determinan representan un tema central en las investigaciones.

Por dichas razones mencionadas hoy en día las universidades cuentan con procesos académicos digitalizados y a nivel de aprendizaje automático se han desarrollado diferentes trabajos de investigación que son mencionados en los antecedentes nacionales e internacionales, utilizando diferentes algoritmos inteligentes, adicionalmente podemos mencionar que según, Castrillón et al. (2020) utilizó el Algoritmo J4; Fernandes et al. (2019) utilizó Minería de datos; Contreras et al. (2020) utilizó Algoritmos de Máquina de Vectores de Soporte y Perceptrón (red neuronal), acordes a los escenarios del ámbito de estudio.

Hasta ahora las diversas investigaciones predictivas basadas en el rendimiento académico han aplicado algoritmos de aprendizaje automático de manera independiente considerando el ámbito de la predicción, tal como se citan en los **antecedentes** del presente trabajo, por tanto, el problema destaca si el algoritmo de aprendizaje automático se relaciona mejor para todos los problemas y si es relevante en la predicción; *por ello el presente trabajo de investigación está enfocado en aplicar diferentes algoritmos de aprendizaje en paralelo a fin de conocer el grado de entrenamiento, aprendizaje y efectividad en los resultados predictivos del ámbito de estudio de la presente investigación.* Asimismo, acorde a la revisión bibliográfica el nuevo conocimiento busca ampliar procedimientos novedosos a fin de magnificar el prestigio y mejora de la calidad educativa en EPISI de UNAM.

En EPISI UNAM aún no ha logrado integrar la productividad educativa a nivel del aprendizaje automático en la gestión académica.

El problema identificado será planteado mediante la aplicación del método analítico que consiste en presentar los síntomas, luego identificar las causas de dichos síntomas,

posteriormente realizar un pronóstico de lo que pasaría si es que no se hace nada respecto a la problemática y finalmente se presenta al control del pronóstico mediante la alternativa de solución de la presente investigación.

El problema de la productividad en EPISI de UNAM se describe de la siguiente manera:

Síntomas de la problemática investigada:

Al respecto se ha determinado lo siguiente:

1. La clasificación de datos y las predicciones de los algoritmos de árboles de decisión son importantes y aplicables al caso, pero no se viene dando en la escuela profesional de ingeniería de sistemas e informática de la UNAM, debido a que no cuenta con una adecuada información debido a que *no se realiza el proceso de inspección, limpieza, transformación y modelación del conjunto de datos* con el objetivo de encontrar información útil que sugiera conclusiones que apoyen a toma de decisiones en las actividades académicas.
2. En la escuela profesional de ingeniería de sistemas e informática de la UNAM en sus procesos no *realizan exploración*, clasificación de datos *para el* modelamiento de probabilidades predictivas del algoritmo naive bayes en las actividades académicas.
3. Asimismo, en la escuela profesional de ingeniería de sistemas e informática de la UNAM en sus procesos académicos no se extraen perspectivas valiosas de los datos, *no realizan sobreajustes en el análisis de datos* del algoritmo random forest que apoyen en la toma de decisiones en las actividades académicas.
4. Se ha determinado que en la escuela profesional de ingeniería de sistemas e informática de la UNAM en sus procesos automatizados no aplican el aprendizaje automático, limitando la ubicación de patrones de comportamiento que permitan desarrollar conocimiento para la toma de decisiones.

Causas de la problemática identificada:

Al respecto se ha determinado lo siguiente:

- a) La falta de ponderación en los algoritmos automatizados de la escuela profesional de ingeniería de sistemas e informática de la UNAM, no ha permitido disponer de políticas ni estrategias algorítmicas para la aplicación del algoritmo de aprendizaje automático de árboles de decisión.
- b) Tampoco se dispone de un marco tecnológico para promover la innovación en la gestión de datos que apoyen a los entrenamientos y modelamientos de los datos en base al algoritmo de aprendizaje naive bayes.



- c) Además, la escuela profesional de ingeniería de sistemas e informática de la UNAM no se cuenta con una estrategia basados en algoritmos ágiles, paralelos, multifuncional tipo random forest que apoyen en la reducción de tiempo y costos.
- d) Asimismo, la escuela profesional de ingeniería de sistemas e informática de la UNAM no se dispone de políticas de innovación tecnológica para la implementación de sistemas de aprendizaje automático.

#### Pronostico

1. Si la escuela profesional de ingeniería de sistemas e informática de la UNAM no aplicara medidas correctivas respecto a los síntomas continuarán las causas y los efectos antes indicados.
2. *Si no se realiza el proceso de inspección, limpia, transformación y modelación del conjunto de datos* no se permitirá encontrar información útil que sugiera conclusiones que apoyen a toma de decisiones en las actividades académicas.
3. No se logrará organizar, concretar con éxito la exploración de datos y sobreajustes para una adecuada clasificación de datos y predicción de los procesos académicos.
4. No se logrará implementar el sistema de aprendizaje automático que ayude a ubicar patrones de comportamiento en el proceso académico.

#### Control del pronóstico de la situación problemática

Para solucionar los síntomas como las causas antes mencionadas en la presente investigación, se propone como control del pronóstico en la productividad educativa mediante los siguientes elementos:

1. *En la aptitud* académica permitirá obtener información referente a la capacidad de los conocimientos adquiridos, motivaciones adquiridas y el desarrollo planificado de sus actividades productivas en la EPISI de UNAM.
2. *En el Medio ambiente* académico, permitirá conocer apoyo del entorno familiar, el desenvolvimiento en el entorno comunitario social, asimismo la influencia de los medios de comunicación en la productividad educativa en EPISI de UNAM.
3. *En referencia a la Instrucción* académica permitirá planificar los estudios, uso de medios tecnológicos y la calidad de instrucción recibida en el proceso de productividad educativa en EPISI de UNAM.
4. Asimismo, el aprendizaje académico permitirá conocer el logro cognitivo obtenido en el proceso de la productividad educativa.

## **Delimitación de la investigación**

### **a) Delimitación espacial**

El presente trabajo de investigación se realizó en la Universidad Nacional de Moquegua en el año académico 2023, ubicado en Ciudad Jardín s/n del distrito de Pacocha, provincia de Ilo, departamento de Moquegua.

### **b) Delimitación social**

La unidad de análisis correspondió a los estudiantes de la Escuela Profesional de Ingeniería de Sistemas e Informática de la Universidad Nacional de Moquegua, periodo 2023.

### **c) Delimitación temporal**

Esta investigación se inició el 01 de marzo y se culminó en diciembre del año 2023.

### **d) Delimitación conceptual**

Se estudiaron las variables:

**a) Productividad educativa:** “Se centra en identificar y comprender el valor atribuido en los conocimientos de aprendizaje adquiridos que contribuyen al éxito de la productividad educativa mejorando la efectividad de los sistemas educativos; propone los componentes de Aptitud, Instrucción, Medio ambiente y Aprendizaje (Figura 4) que se combinan e influyen entre sí, a fin de determinar el éxito del rendimiento final del estudiante” (Mazana et al., 2019), (Peconcillo Jr et al., 2020).

**b) Algoritmos de aprendizaje automático:** “Es parte de la ciencia de datos, que mediante el uso de métodos estadísticos, los algoritmos son programados, entrenados a fin de realizar clasificaciones o predicciones a fin de conocer información importante, en donde dichos conocimientos permitirán la toma de decisiones dentro de las aplicaciones y las organizaciones empresariales, a fin de generar impacto en su cadena de valores (IBM Cloud Education, 2020), esto significa que los algoritmos se vuelven inteligentes y más en el tiempo, permitiendo establecer efectos positivos de fortalecimiento institucional” (Sharma et al., 2021).

## **Problemas de investigación**

### **Problema principal**

- ¿Cuál es el nivel de relación existe entre la productividad educativa y los algoritmos de aprendizaje en estudiantes EPISI de UNAM, 2023?

### **Problemas específicos**

- ¿Cuál es el nivel de relación que existe entre la dimensión aptitud y los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023?
- ¿Cuál es el nivel de relación que existe entre la dimensión instrucción y los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023?
- ¿Cuál es el nivel de relación que existe entre la dimensión medio ambiente y los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023?
- ¿Cuál es el nivel de relación que existe entre la dimensión aprendizaje y los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023?

### **Objetivos de la investigación**

#### **Objetivo general**

- Determinar el nivel de relación que existe entre la productividad educativa y los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023

#### **Objetivos específicos**

- Determinar el nivel de relación que existe entre la dimensión aptitud y los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023
- Establecer el nivel de relación que existe entre la dimensión instrucción y los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023
- Especificar el nivel de relación que existe entre la dimensión medio ambiente y los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023
- Analizar el nivel de relación que existe entre la dimensión aprendizaje y los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023

### **Justificación e importancia de la investigación.**

Según Hernández-Sampieri R. (2018) en la investigación, la justificación de una investigación implica detallar las razones que respaldan su realización. Esto significa que justificar un estudio consiste en presentar los motivos que demuestran la relevancia de llevar a cabo dicho análisis, fundamentándose en los objetivos y las preguntas planteadas en la investigación actual.

#### **a) Justificación teórica**

Considerando la existencia del enfoque de modelado de datos y considerando la existencia de algoritmos iniciales en donde son interpretados; machine learning automatiza el proceso de

datos mediante la producción de modelos tanto en los datos actuales, como en los futuros, proporcionando una aceleración sin precedentes concernientes a la productividad educativa basado en el rendimiento académico de EPISI de la Universidad Nacional de Moquegua. El marco teórico concerniente a la productividad educativa esta referenciado por el modelo de productividad educativa propuesto por Herbert J. Walberg, basado en la teoría del aprendizaje cognitivo; en referencia a los algoritmos de aprendizaje automático se consideran Arboles de decisión, Naive Bayes y Random Forest, a fin de realizar un análisis predictivo mediante el aprendizaje de los datos de reciente recogida sin necesidad de intervención humana y sin hacer falta de una programación explícita.

#### **b) Justificación práctica**

Conscientes de la importancia que va adquiriendo el uso correcto de los sistemas inteligentes basados en modelos de aprendizaje predictivo en tecnologías de la información enfocados al rendimiento académico en la Universidad Nacional de Moquegua y los indicadores edad, genero, lugar de procedencia, centro laboral, tipo de preparación, promedio final, modalidad de ingreso, cursos matriculados, escuela profesional, tipo de matrícula, porcentaje de predicción acertada y el uso de la eficiencia de los algoritmos de aprendizaje automática y que esta justificación metodológica ayudara en la mejora de la productividad académica en la Universidad Nacional de Moquegua.

El presente trabajo de investigación tuvo como objetivo observar la relevancia del uso del uso de la eficiencia de los algoritmos de aprendizaje automático para el sistema académico basado en la productividad educativa de EPISI de la UNAM.

#### **c) Justificación metodológica**

Acorde a los objetivos de la investigación, se llevó a cabo un proceso metodológico estructurado y sistemático, empleando técnicas de investigación cuantitativa enfocadas en el análisis y la síntesis respecto al uso de algoritmos de aprendizaje automático para la predicción de la productividad educativa en base al rendimiento académico en la escuela profesional de ingeniería de sistemas e informática de la Universidad Nacional de Moquegua., determinando en ambas variables los procedimientos para la jerarquización de los factores descriptivos y explicativos.

#### **d) Justificación social**

Las universidades hoy en día se involucran en actividades sociales mediante la divulgación de los conocimientos, para ello es importante la productividad académica a fin de elevar los conocimientos junto a la innovación y transferencia tecnológica, basado en la sociedad del conocimiento. El protagonismo de las universidades peruanas en la creación, transmisión y

aplicación de conocimiento hace prescindible el uso de algoritmos de aprendizaje para predecir los rendimientos académicos y fortalecer los conocimientos adquiridos en la UNAM.

### **Importancia**

La importancia de aplicar nuevas técnicas inteligentes en el presente trabajo de investigación establecerá nuevos horizontes para realizar estudios sustantivos que presenten situaciones similares planteadas en la presente investigación, asimismo el marco referencial contribuirá a la implementación de políticas educativas relacionadas al mejoramiento de la productividad educativa basado en logros de aprendizaje en un marco de equidad y justicia. También considerar que las nuevas técnicas de aprendizaje automático brindarán información valiosa para mejorar las prácticas pedagógicas y establecer las estrategias de aprendizaje que afectan el rendimiento académico.

### **Factibilidad de la investigación**

Se dispuso de recursos como tiempo, asesoría continua y acceso a información, aunque en algunos casos esta fue limitada. También se contaron con recursos económicos restringidos y una actitud personal comprometida para llevar a cabo la investigación de manera responsable y completa, con el fin de graduarme. Los recursos mencionados se utilizaron para evaluar la viabilidad de la investigación, mediante el cumplimiento del objetivo específico.

### **Limitaciones del estudio**

El desarrollo tuvo algunas limitaciones referentes a la obtención de los datos para el presente estudio, para ello se consideró estudiantes regulares e irregulares matriculados en la escuela profesional de ingeniería de sistemas de UNAM, con un promedio de edad 20 hasta los 26 años de edad aproximadamente concernientes a EPISI.

Es importante mencionar que el presente estudio estuvo orientado a obtener información relevante conforme al modelo de productividad educativa propuesto por Herbert J. Walberg en las dimensiones de aptitud, instrucción, medio ambiente y aprendizaje, ampliado con indicadores correspondiente a fin de fortalecer las variables de estudio, asimismo los algoritmos en estudio desarrollaran el análisis, clasificación y aprendizaje respectivo afianzando el propósito del presente trabajo de investigación.

El acceso a los datos a la población objetivo puede resultar difícil para la investigación, para ello los cuestionarios se aplicaron en forma directa en las inmediaciones de la Escuela Profesional de Ingeniería de Sistema e Informática de la Universidad Nacional de Moquegua.



Las limitaciones temporales no pudieron afectar la cantidad de datos recopilados, para ello se tuvo que planificar cuidadosamente el cronograma de investigación, establecer metas realistas y utilizar métodos eficientes de recolección de datos, junto a la flexibilización y la adaptación ante imprevistos.

## Referencias

- Brew, E. A. y Nketiah, B. y Koranteng, R. (2021). A Literature Review of Academic Performance, an Insight into Factors and their Influences on Academic Outcomes of Students at Senior High Schools. *OALib*, 08(06), 1–14. <https://doi.org/10.4236/oalib.1107423>
- Castrillón, O. D. y Sarache, W. y Ruiz-Herrera, S. (2020). Predicción del rendimiento académico por medio de técnicas de inteligencia artificial. *Formación Universitaria*, 13(1), 93–102. <https://doi.org/10.4067/S0718-50062020000100093>
- Contreras, L. E. y Fuentes, H. J. y Rodríguez, J. I. (2020). Predicción del rendimiento académico como indicador de éxito/fracaso de los estudiantes de ingeniería, mediante aprendizaje automático. *Formación Universitaria*, 13(5), 233–246. <https://doi.org/10.4067/S0718-50062020000500233>
- Edgard, T. W. y Manz, D. O. (2017). *Research Methods for Cyber Security* (1ra ed.). Elsevier Inc.
- Fernandes, E. y Holanda, M. y Victorino, M. y Borges, V. y Carvalho, R. y Erven, G. van. (2019). Educational data mining: Predictive analysis of academic performance of public school students in the capital of Brazil. *Journal of Business Research*, 94, 335–343. <https://doi.org/10.1016/j.jbusres.2018.02.012>
- Hawkins, C. y Bailey, L. E. (2020). A New Data Landscape: IR's Role in Academic Analytics. *New Directions for Institutional Research*, 2020(185–186), 87–103. <https://doi.org/10.1002/ir.20331>
- Hernández-Sampieri R. (2018). *Metodología de la investigación* (1ra Ed.).
- IBM Cloud Education. (2020). *Machine Learning This introduction to machine learning provides an overview of its history, important definitions, applications, and concerns within businesses today*. <https://www.ibm.com/cloud/learn/machine-learning>
- Mazana, M. Y. y Montero, C. S. y Casmir, R. O. (2019). Investigating Students' Attitude towards Learning Mathematics. *International Electronic Journal of Mathematics Education*, 14(1). <https://doi.org/10.29333/iejme/3997>

- McGinnis, D. (2020). *¿Qué es la Cuarta Revolución Industrial? | Fuerza de ventas*.  
<https://www.salesforce.com/blog/what-is-the-fourth-industrial-revolution-4ir/>
- Medina N. y Fereira J. y Marzol R. (2018). Factores personales que inciden en el bajo rendimiento académico de los estudiantes de geometría. *Revista de Estudios Interdisciplinarios En Ciencias Sociales*.
- Peconcillo Jr, L. B. y D. Peteros, E. y O. Mamites, I. y T. Sanchez, D. y L. Tenerife, J. J. y L. Suson, R. (2020). Structuring Determinants to Level Up Students Performance. *International Journal of Education and Practice*, 8(4), 638–651.  
<https://doi.org/10.18488/journal.61.2020.84.638.651>
- Ray, S. (2019). A Quick Review of Machine Learning Algorithms. *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, 35–39. <https://doi.org/10.1109/COMITCon.2019.8862451>
- Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, 2(3), 160. <https://doi.org/10.1007/s42979-021-00592-x>
- Sharma, N. y Sharma, R. y Jindal, N. (2021). Machine Learning and Deep Learning Applications-A Vision. *Global Transitions Proceedings*, 2(1), 24–28.  
<https://doi.org/10.1016/j.gltp.2021.01.004>
- Triwiyanto T. y Suyanto y Prasajo L.D. y Wardana Y. (2020). Factors Affecting Educational Productivity at Private Elementary Schools in Indonesia. *Advances in Social Science, Education and Humanities Research*, 487, 318–323.

## CAPÍTULO 2

### Marco filosófico

### Philosophical framework

DOI: <https://doi.org/10.7112/53tgaw73>

#### Resumen

El capítulo desarrolla el marco filosófico de la investigación desde los enfoques ontológico y epistemológico, sustentado en el paradigma positivista. Se fundamenta la productividad educativa a partir del modelo de Herbert J. Walberg, orientado a la medición objetiva del rendimiento académico, la eficiencia y la mejora continua. Asimismo, se establece la base ontológica y epistemológica de los algoritmos de aprendizaje automático, destacando su carácter empírico, probabilístico e instrumental para la predicción académica. Ambos enfoques sustentan científicamente el estudio de la relación entre productividad educativa y aprendizaje automático.

**Palabras clave:** Marco filosófico, ontología, epistemología, productividad educativa, aprendizaje automático.

#### Abstract

This chapter develops the philosophical framework of the research from ontological and epistemological perspectives, grounded in the positivist paradigm. Educational productivity is supported through Herbert J. Walberg's model, which emphasizes the objective measurement of academic performance, efficiency, and continuous improvement. Likewise, the ontological and epistemological foundations of machine learning algorithms are established, highlighting their empirical, probabilistic, and instrumental nature for academic prediction. Both approaches provide scientific support for the study of the relationship between educational productivity and machine learning.

**Keywords:** Philosophical framework, ontology, epistemology, educational productivity, machine learning.

Existe un razonamiento preliminar basado en el conocimiento y percepciones cuando se empieza a filosofar propios de los conocimientos adquiridos, según Saunders M. et al. (2019) la filosofía de la investigación es el desarrollo racional de los supuestos conocimientos y su

naturaleza; asimismo, se basa en *supuestos*; en donde dicha suposición es la declaración preliminar del razonamiento (Hitchcock, Hughes, 1989, como se citó en Žukauskas et al., 2018), lo cual significa que pueden existir diferentes supuestos sobre la naturaleza de la verdad, el conocimiento y su adquisición.

La filosofía de investigación se refiere a la manera adecuada de recolectar, analizar y aplicar datos sobre un fenómeno. El concepto de epistemología (lo que se considera verdadero) contrasta con la doxología (lo que se cree que es verdadero) e incluye las distintas filosofías relacionadas con los enfoques de investigación. El propósito de la ciencia, entonces, es el proceso de transformar las cosas creídas en cosas conocidas: doxa en episteme. Asimismo, según, Orlikowski W. y Baroudi J.(1991) se determinan tres paradigmas relevantes en Sistemas de Información: positivista, interpretativo y crítico.

Según Biblioteca Nacional de Chile (2021), la filosofía positivista defendía que el único conocimiento genuino es el científico, sustentado en el modelo empírico-formal, el cual es un legado de la racionalidad moderna. Este enfoque se basa en las contribuciones de la cosmovisión de Galileo, la Ciencia Nueva, y el racionalismo cartesiano propuesto por René Descartes en donde pretendía alcanzar la verdad más allá de toda duda (Hernandez-Manxilla J.M., 2011), por tanto el presente trabajo de investigación, está enmarcado dentro del paradigma positivista, porque busca medir y encontrar leyes, basados en la prueba y ser validados y las relaciones existentes en la estadística, matemática, cuyo objetivo es encontrar y/o medir la relación entre las variables involucradas dentro del escenario o ámbito del presente trabajo de investigación.

### **Fundamentación ontológica**

La fundamentación ontológica se centra en la descripción de la estructura semántica de la información basado en el razonamiento del concepto ontológico del dominio (Padilla-Cuevas et al., 2021), para nuestro caso es la estructura semántica de las variables involucradas de la presente investigación, para ello la ontología es una herramienta necesaria para realizar exploraciones y análisis a fin de caracterizar la investigación (Salatino et al., 2020).

Para ello el positivismo sostiene una postura de la existencia de una realidad tangible que puede ser entendida, identificada y medida. (Park et al., 2020).

La ontología representa un conjunto de conceptos compartidos, acordado y sintonizado en donde explican un determinado problema, según McDaniel y Storey (2020) una de las principales ventajas de utilizar una ontología de dominio es su capacidad para definir un modelo semántico de los datos combinados con el conocimiento del dominio asociado, es decir

permiten estar integrados e intercambien información. Significa adicionalmente que la ontología de dominio debe garantizar una comunicación significativa, sintonizando los diferentes tipos de conocimiento semántico a fin de optimizar los datos en información valedera.

La *productividad educativa* basada en el rendimiento académico ha generado numerosos estudios teóricos y empíricos, a fin de medir el rendimiento de los estudiantes mediante pruebas estandarizadas, según Universidad de Lehigh (2013) Horace Mann en 1845 elabora exámenes escritos en Boston a fin de establecer el rendimiento académico y la calidad de educación que estaban recibiendo, dicho suceso creó consecuencias negativas publicadas en el The New York Times, posteriormente provocó una maduración interesante respecto a establecer una taxonomía significativa a través de Bloom a fin de establecer objetivos de aprendizaje, luego en 1963 John B. Carroll propone un modelo para el aprendizaje escolar, basado en el tiempo del aprendizaje y el tiempo realmente necesario para el aprendizaje, luego en 1984 Herbert J. Walberg propone el Modelo de productividad educativa, donde establece factores que influyen en el rendimiento académico; por tanto considerando lo mencionado el ámbito ontológico estará centrado en la propuesta taxonómica acorde a las dimensiones del Modelo de Walberg, cuya construcción pretende establecer un mejor conocimiento del ámbito del rendimiento académico.

Por tanto, la **fundamentación ontológica** de la productividad educativa propuesta por Herbert J. Walberg **se encuentra arraigada en su perspectiva filosófica sobre la naturaleza de la realidad y el ser**. La ontología se refiere al estudio de la naturaleza del ser y la realidad, y en el caso de Walberg, su enfoque ontológico se relaciona con la búsqueda de comprender la esencia de la educación y cómo esta puede ser más efectiva.

Herbert J. Walberg, un destacado psicólogo y educador, ha abordado la cuestión de la productividad educativa desde una perspectiva que integra elementos psicológicos, pedagógicos y sociales. Aunque no existe una formulación ontológica explícita en sus obras, se pueden identificar algunos principios fundamentales que subyacen en su enfoque:

- **Énfasis en los resultados medibles:** Walberg aboga por una perspectiva orientada a resultados en educación, donde la efectividad se evalúa a través de medidas cuantificables. Esto implica una creencia ontológica en la capacidad de medir y evaluar los resultados educativos de manera objetiva.
- **Importancia de los factores externos y contextuales:** Su enfoque reconoce la influencia de factores externos y contextuales en el proceso educativo. Esto implica una



ontología que considera la interconexión y la interdependencia de diversos elementos que influyen en la eficacia educativa.

- **Énfasis en la eficiencia y la mejora continua:** La productividad educativa según Walberg está relacionada con la eficiencia y la mejora continua. Esto sugiere una ontología que ve la educación como un proceso dinámico y en constante evolución, donde la optimización y la eficiencia son valores fundamentales.
- **Valorización de la investigación empírica:** Walberg aborda la productividad educativa desde una perspectiva basada en la evidencia empírica. Su enfoque ontológico implica la creencia en la capacidad de la investigación científica para proporcionar conocimientos válidos sobre cómo mejorar la educación.

La fundamentación ontológica de la productividad educativa propuesta por Herbert J. Walberg puede caracterizarse por una orientación hacia la medición objetiva de resultados, la consideración de dimensiones contextuales, la importancia de la eficiencia y la mejora continua, y la valorización de la investigación empírica en el campo educativo. Estos principios ontológicos influyen en su enfoque metodológico y en su visión sobre cómo lograr una educación más productiva y efectiva.

En relación a la variable algoritmos de aprendizaje automático, acorde a lo mencionado por Mahesh (2018), señala que esta área se enfoca en el estudio científico de algoritmos y modelos matemáticos empleados en los sistemas automatizados para llevar a cabo tareas específicas, sin necesidad de desarrollar una programación explícita.

por tanto, el aprendizaje automático va a utilizar algoritmos que razonan a partir de datos proporcionados externamente, producir hipótesis generales y que luego realizar predicciones sobre instancias futuras (Osisanwo et al., 2017); la ontología estará centrado en los dominios de clasificación de los algoritmos de aprendizaje automático: **Árboles de decisión**; cuyo aprendizaje consiste en un conjunto de datos que serán divididos recursivamente en subconjuntos y así sucesivamente irán mejorando las divisiones de los subconjuntos según la pureza obtenida por el subconjunto. La pureza consiste en la obteniendo puntuaciones en proporción a las clases individuales en una mezcla de etiquetas de la clase. Cuanto mayor sea la proporción en una de las clases, más pura será la colección (Yeturu, 2020), **Naive Bayes**, el aprendizaje estará basado en el teorema de Bayes, cuya clasificación se basa en la clasificación probabilística condicional simple, creando árboles en función de la probabilidad de lo que sucederá, asimismo mantiene una fuerte independencia de los valores de los atributos (Sarker, 2021b); **Random Forest**, el algoritmo se basa en la clasificación de entrenamiento de varios árboles de decisión en paralelo, esto significa que el entrenamiento se realiza

paralelamente por árbol de manera independiente según sus propios subconjuntos y características disponibles (Misra y Li, 2020), cuyos aportes serán significativos para la toma de decisiones basadas en el rendimiento académico.

Por ello, la **fundamentación ontológica de los algoritmos de aprendizaje automático (machine learning)** se basa en una comprensión particular de la realidad y del proceso de aprendizaje. Aunque el aprendizaje automático se enmarca principalmente en un contexto informático y matemático, su fundamento ontológico implica ciertos supuestos filosóficos sobre la naturaleza del conocimiento y la capacidad de las máquinas para aprender, mencionamos los aspectos clave de la fundamentación ontológica de los algoritmos de aprendizaje automático:

- **Realismo instrumental:** La ontología del aprendizaje automático a menudo adopta una perspectiva realista instrumental. En este enfoque, se considera que los modelos generados por los algoritmos son herramientas efectivas para hacer predicciones y tomar decisiones, pero no necesariamente reflejan una comprensión profunda de la realidad subyacente.
- **Construccionismo:** Desde una perspectiva ontológica, el aprendizaje automático puede tener raíces en el construccionismo, que sostiene que el conocimiento es construido activamente por la mente en lugar de ser una representación pasiva de la realidad. En el contexto del aprendizaje automático, los algoritmos construyen modelos a partir de datos, y estos modelos son construcciones que permiten hacer predicciones útiles.
- **Epistemología probabilística:** La ontología de los algoritmos de aprendizaje automático a menudo asume una epistemología probabilística. Esto implica que la incertidumbre y la variabilidad son características intrínsecas del conocimiento y de la realidad. Los modelos de aprendizaje automático suelen expresar predicciones en términos de probabilidades, reconociendo la naturaleza probabilística de la inferencia.
- **Reduccionismo y abstracción:** Desde un punto de vista ontológico, los algoritmos de aprendizaje automático adoptan una perspectiva reduccionista y abstracta. Los modelos generados por estos algoritmos simplifican la complejidad de los datos y los patrones subyacentes para facilitar la comprensión y la toma de decisiones.
- **Supervisión y aprendizaje inductivo:** La ontología del aprendizaje automático a menudo se alinea con una epistemología inductiva, donde se infieren patrones generales a partir de ejemplos específicos. La supervisión, en la que los modelos se

entrenan con datos etiquetados, es un componente fundamental, reflejando la idea de aprender a partir de casos concretos.

La fundamentación ontológica de los algoritmos de aprendizaje automático implica una perspectiva instrumental y construccionista, una epistemología probabilística que reconoce la incertidumbre inherente, una orientación hacia la reducción y abstracción de la complejidad, y un enfoque inductivo centrado en el aprendizaje a partir de ejemplos específicos. Estos principios filosóficos influyen en cómo se concibe y aplica el aprendizaje automático para abordar problemas en diversos campos.

### **Fundamentación epistemológica**

Según Al-Ababneh M. (2020) la epistemología proporciona una base filosófica de los posibles tipos de conocimientos adecuados que aseguran la legitimidad de la investigación., es decir es la aceptación del conocimiento en el campo de estudio (Saunders et al., 2009, como se citó en Al-Ababneh M., 2020).

Epistemología es la teoría del conocimiento, en donde se destaca las explicaciones de los fenómenos, los mismos que son validados con evidencias (Giacomini, 2010, como se citó en Hiller, 2016); otra versión define a la epistemología como el conocimiento adquirido que implica el cómo sabemos lo que sabemos (Crotty, 1998 como se citó en Al-Ababneh, 2020), desde el ámbito filosófico se examina la naturaleza, las limitaciones y la justificación del conocimiento (Williams 2001 como se citó en Hathcoat et al., 2019), por tanto la Epistemología se basa en la aceptación del conocimiento según el campo de estudio (Saunders et al., 2029); por ello el presente trabajo destaca variables en el ámbito académico y algoritmos de aprendizaje automático, con la finalidad de establecer aportes en el escenario descrito.

Hoy en día en las instituciones educativas se evalúan el proceso enseñanza aprendizaje a fin de cuantificar el logro obtenido motivando la recopilación de datos del proceso a fin de conocer el rendimiento académico del estudiante; según (Shadish y Luellen como se citó en Alkin y King, 2016), la historia de la evaluación es tan antigua como la historia humana, básicamente consiste: a) identificar un problema, b) implementar alternativas para reducir síntomas, c) evaluar alternativas y d) adoptar el resultado de dicha evaluación; esto significa que las personas se involucran en actividades de evaluación a fin de animar su uso; asimismo los referenciados formulan evidencias que remonta a miles de años basado en la evaluación de dieta citado en el capítulo 1 del libro de Daniel de la Biblia y la evaluación del personal en China, hace más de 4000 años, dichos ejemplos son una muestra histórica de la evaluación; incluso la definición de evaluación según Scriven (1967), resalta mayor significancia basada en el "valor"

y "utilidad" de la evaluación con la finalidad de ser utilizadas en un contexto dado en un momento dado.

Horace Mann considerado “el padre de las pruebas estandarizadas en los EE. UU” propuso en 1843 a la Junta de Educación del Estado de Massachusetts la elaboración de exámenes escritos a fin de conocer la calidad educativa de los estudiantes de las escuelas de Boston cuyas evaluaciones los maestros desconocían y cuyo resultado generó un impacto negativo en la sociedad con consecuencias a niveles de despidos; posteriormente dichos antecedentes motivaron realizar aportes significativos a fin de mejorar la analítica en el ámbito académico; para ello Benjamin S. Bloom en 1956 y un grupo de psicólogos educativos desarrollaron una jerarquía de objetivos educativos, basado en los resultados académicos, luego en 1963 John B. Carroll propone un modelo para el aprendizaje escolar, basado en el tiempo del aprendizaje y el tiempo realmente necesario para el aprendizaje, en 1984 Herbert J. Walberg propone el Modelo de productividad educativa, basado en factores del rendimiento académico; el mencionado hecho de medir los atributos de los factores relacionados a la productividad académica y a lo mencionado por Scriven acerca del "valor" y "utilidad" en un contexto dado, en un momento dado establece la epistemología de la variable según el ámbito de estudio del presente trabajo.

Por tanto, **la fundamentación epistemológica de la productividad educativa** propuesta por Herbert J. Walberg se basa en su enfoque en la aplicación de métodos científicos y la búsqueda de conocimientos objetivos y medibles en el campo educativo. Walberg ha abogado por un enfoque basado en la evidencia y la medición cuantitativa para evaluar y mejorar la efectividad de los sistemas educativos. A continuación, se destacan algunos aspectos clave de la fundamentación epistemológica de la productividad educativa según Walberg:

- **Empirismo y positivismo:** La epistemología de Walberg refleja tendencias empiristas y positivistas, que enfatizan la importancia de la observación, la medición y la verificación empírica en la obtención del conocimiento. Este enfoque sugiere que el conocimiento en educación puede obtenerse de manera más confiable a través de métodos científicos objetivos.
- **Valoración de la investigación cuantitativa:** Walberg ha abogado por la investigación cuantitativa como un medio para obtener datos medibles y objetivos sobre la productividad educativa. Su epistemología sugiere que la recopilación sistemática y el análisis cuantitativo de datos son esenciales para comprender y mejorar la eficacia de los métodos educativos.

- **Énfasis en resultados medibles:** La epistemología de Walberg se alinea con la idea de que los resultados educativos medibles son indicadores válidos de la eficacia de los métodos de enseñanza. Este enfoque implica que el conocimiento sobre la productividad educativa se construye a través de la evaluación objetiva de los resultados obtenidos por los estudiantes.
- **Utilización de métodos de investigación experimental:** Walberg ha respaldado el uso de métodos de investigación experimental en educación, incluyendo estudios de control y diseño experimental. Este enfoque refleja una epistemología que valora la identificación de relaciones causales y la aplicación de intervenciones controladas para evaluar su impacto.
- **Enfoque en la medición del rendimiento:** La fundamentación epistemológica de Walberg destaca la importancia de medir el rendimiento estudiantil como una forma de evaluar la eficacia de los programas educativos. Esto sugiere una epistemología que considera que la medición objetiva es esencial para obtener conocimientos significativos en el ámbito educativo.

Por tanto, la fundamentación epistemológica de la productividad educativa propuesta por Herbert J. Walberg se caracteriza por un enfoque empirista y positivista, con un énfasis en la investigación cuantitativa, la medición objetiva de resultados y la aplicación de métodos experimentales para avanzar en el conocimiento sobre la eficacia de las prácticas educativas.

Respecto a los algoritmos de aprendizaje automático, según De Oliveira Chagas (2019) en 1943 el neurocientífico Warren McCulloch y el lógico Walter Pitts, idearon por primera vez una red neuronal simple en el artículo “A Logical Calculus of the Ideas Immanent in Nervous Activity”, el mismo mapeaba el proceso de toma de decisiones en la cognición humana y las redes neuronales; dicho aporte fue seguido por la “era de las ideas”, entre los años 1950 y 1960, en donde se propusieron varios modelos que podía hacer realidad a la Inteligencia Artificial, en dicho periodo Alan Turing diseñó la prueba de Turing en 1950 (Gopani, 2022), a fin de determinar si una máquina podría emular un comportamiento inteligente similar al ser humano, luego en 1956, Marvin Minsky y John McCarthy realizaron la "Conferencia que lo inició todo" en Dartmouth College, en donde se estableció los principios fundamentales de la Inteligencia Artificial, como describe en el blog de Harvard (Anyoha, 2017). En 1951 se desarrolló el algoritmo del vecino más cercano por Evelyn Fix y Joseph Hodges y luego ampliado en 1967 por Thomas Cover y Peter E. Hart (Auffarth, 2021), en 1952 Arthur Samuel en IBM desarrolló el primer programa para PC de juego de damas, utilizando el algoritmo de poda alfa-beta, un algoritmo de búsqueda que disminuye nodos a través del algoritmo minimax

en los árboles de búsqueda, en 1958 Frank Rosenblatt desarrollo el “perceptrón” con la finalidad de superar las limitaciones del modelo neuronal de McCulloch-Pitts dando inicio el concepto de entrenamiento supervisado (Rosenblatt, 1958), en 1981, Gerald Dejong descubrió aprendizaje basado en la explicación , basado en un algoritmo informático, que incluía la explicación y la generalización de datos, en 1985 desarrollan NetTalk por Terry Sejnowski, un algoritmo, basado en redes neuronales es capaz de pronunciar palabras como un bebé, basado en texto y transcripción fonética coincidente como entrada, en 1989, Christopher Watkins desarrolló un algoritmo Q-learning mejorando las prácticas del aprendizaje por refuerzo, en 1990 los métodos estadísticos se difundieron para para los algoritmos, debido a que las redes neuronales parecían menos explicables y necesitaban recursos computacionales más exigentes; dichos métodos incluían el algoritmo de máquinas con vectores de soporte y el algoritmo bosque aleatorio introducidos en 1995, luego surge la trascendencia muy importante para la inteligencia artificial en 1997 de IBM Deep Blue logra vencer al campeón mundial de ajedrez Garri Kasparov (Pastor, 1997).

En el año 2009 Fei-Fei Li, profesora de informática de la Universidad de Stanford introduce las redes convolucionales en donde creó un gran conjunto de datos que reflejaba el mundo real, posteriormente se convirtió en la base de Alex Krizhevsky, quien creó la primera CNN (redes neuronales convolucionales), AlexNet, en 2012; en 2011, Watson de IBM presentó su máquina que puede categorizar objetos como un gato, el mismo fue mejorado y puesta en prueba en YouTube en 2012, en 2013 desarrollaron algoritmos de Word2vec con redes neuronales para aprender asociaciones de palabras y luego se convirtieron en la base de grandes modelos de lenguaje, en 2014, Facebook desarrolló DeepFace , un algoritmo de reconocimiento de rostros humanos. En 2014 también vio la creación de redes antagónicas generativas (GAN) por parte de Ian Goodfellow.

En 2016, en aprendizaje automático se desarrolló Deepmind AlphaGo, que venció al campeón mundial en el juego de mesa chino, Go; en 2017, AlphaGo y sus sucesores vencieron a varios campeones de Go, Ajedrez y Shogi, asimismo en 2017, Waymo comenzó a probar sus minivans autónomas; en 2018 la compañía de IA Deepmind desarrolla el programa de IA AlphaFold logrando predecir la estructura de las proteínas; en 2020, Facebook AI Research introdujo el aprendizaje basado en creencias recursivas, o ReBeL, un algoritmo general de RL+Search con la finalidad de funcionar en todos los juegos de dos jugadores (Gopani, 2022). El escenario de la inteligencia artificial es indudablemente el campo multidisciplinario que ha desarrollado grandes avances del conocimiento, es por ello que la epistemología de la variable algoritmos de aprendizaje automático busca dar a conocer nuevos aportes en el campo

educativo considerando modelos matemáticos de clasificación a fin de predecir el rendimiento académico en EPISI UNAM.

La **fundamentación epistemológica de los algoritmos de aprendizaje automático** (AA) se basa en una perspectiva principalmente empirista y pragmática. Aunque los algoritmos de aprendizaje automático pertenecen al ámbito de la inteligencia artificial y la informática, su epistemología refleja cómo se adquiere el conocimiento y cómo se interpreta la información.

A continuación, describimos los elementos clave de la fundamentación epistemológica de los algoritmos de aprendizaje automático:

- **Empirismo y pragmatismo:** La epistemología de los algoritmos de aprendizaje automático se alinea con el empirismo al valorar la experiencia y la observación. La eficacia de estos algoritmos se evalúa a través de su capacidad para aprender de datos empíricos y realizar predicciones útiles en entornos específicos. Además, adopta un enfoque pragmático al enfocarse en soluciones prácticas y resultados observables.
- **Inductivismo y generalización:** La epistemología de los Algoritmos de aprendizaje automático se basa en el inductivismo, donde se infieren patrones generales a partir de ejemplos específicos. Estos algoritmos buscan generalizar patrones a partir de conjuntos de datos para hacer predicciones en situaciones nuevas. La capacidad de generalización es esencial para la aplicabilidad de los modelos generados.
- **Probabilismo e incertidumbre:** Dado que los algoritmos de aprendizaje automático a menudo trabajan con datos ruidosos y situaciones complejas, su epistemología incorpora la noción de probabilismo. Los modelos generados expresan predicciones en términos de probabilidades, reconociendo la incertidumbre inherente en la inferencia a partir de datos limitados o ruidosos.
- **Instrumentalismo:** La epistemología de los AA tiende a adoptar una perspectiva instrumentalista, donde los modelos generados son considerados como herramientas útiles para realizar tareas específicas en lugar de representaciones precisas de la realidad. Estos modelos se evalúan por su utilidad y capacidad para realizar predicciones precisas, más que por su fidelidad a una verdad objetiva.
- **Rechazo de la verdad absoluta:** La epistemología de los algoritmos de aprendizaje automático tiende a ser pragmática y no busca necesariamente una verdad absoluta. Se centra en la utilidad y la eficacia en contextos específicos, reconociendo que diferentes modelos pueden ser más apropiados para diferentes situaciones.

Por tanto, la fundamentación epistemológica de los algoritmos de aprendizaje automático se basa en el empirismo, el inductivismo, el probabilismo y el instrumentalismo. Estos principios reflejan la naturaleza práctica y utilitaria de los algoritmos de aprendizaje automático, donde el conocimiento se adquiere a través de la experiencia, la observación y la inferencia a partir de datos empíricos.

## Referencias

- Al-Ababneh M. (2020). Linking Ontology, Epistemology and Research Methodology. *Science & Philosophy*.
- Al-Ababneh, M. M. (2020). Linking Ontology, Epistemology and Research Methodology. *Science & Philosophy*, 75–91.
- Alkin, M. C. y King, J. A. (2016). The Historical Development of Evaluation Use. *American Journal of Evaluation*, 37(4), 568–579. <https://doi.org/10.1177/1098214016665164>
- Anyoha, R. (2017). *La historia de la inteligencia artificial - Ciencia en las noticias*. <https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/>
- Auffarth, B. (2021). *Machine Learning for Time-Series with Python*. [https://books.google.com.pe/books?id=a7tLEAAQBAJ&pg=PA99&lpg=PA99&dq=author+of+nearest+neighbors+algorithm++Evelyn+Fix+y+Joseph+Hodges&source=bl&ots=D19ck8A07W&sig=ACfU3UovGNfCvt\\_kI7tpjETbQgZOoGuyAA&hl=es&sa=X&ved=2ahUKewiTt474wcj4AhVVAdQKHUSABNUQ6AF6BAgoEAM#v=onepage&q=author%20of%20nearest%20neighbors%20algorithm%20%20Evelyn%20Fix%20y%20Joseph%20Hodges&f=false](https://books.google.com.pe/books?id=a7tLEAAQBAJ&pg=PA99&lpg=PA99&dq=author+of+nearest+neighbors+algorithm++Evelyn+Fix+y+Joseph+Hodges&source=bl&ots=D19ck8A07W&sig=ACfU3UovGNfCvt_kI7tpjETbQgZOoGuyAA&hl=es&sa=X&ved=2ahUKewiTt474wcj4AhVVAdQKHUSABNUQ6AF6BAgoEAM#v=onepage&q=author%20of%20nearest%20neighbors%20algorithm%20%20Evelyn%20Fix%20y%20Joseph%20Hodges&f=false)
- Biblioteca Nacional de Chile. (2021). *La Filosofía Positivista*. <http://www.memoriachilena.gob.cl/602/w3-article-93966.html>
- de Oliveira Chagas, E. T. (2019). Deep Learning and its applications today. *Revista Científica Multidisciplinar Núcleo Do Conhecimento*, 04(05), 05–26. <https://doi.org/10.32749/nucleodoconhecimento.com.br/business-administration/deep-learning-2>
- Gopani, A. (2022). *La historia de los algoritmos de aprendizaje automático*. <https://analyticsindiamag.com/the-history-of-machine-learning-algorithms/>
- Hathcoat, J. D. y Meixner, C. y Nicholas, M. C. (2019). Ontology and Epistemology. In *Handbook of Research Methods in Health Social Sciences* (pp. 99–116). Springer Singapore. [https://doi.org/10.1007/978-981-10-5251-4\\_56](https://doi.org/10.1007/978-981-10-5251-4_56)



- Hernandez-Manxilla J.M. (2011). El racionalismo cartesiano y su particular conquista de la subjetividad en el mundo moderno. *FRENIA*.
- Hiller, J. (2016). *Epistemological foundations of objectivist and interpretivist research* (pp. 99–127).
- Mahesh, B. (2018). Machine Learning Algorithms - A Review. *International Journal of Science and Research*, 381–386.
- McDaniel, M. y Storey, V. C. (2020). Evaluating Domain Ontologies. *ACM Computing Surveys*, 52(4), 1–44. <https://doi.org/10.1145/3329124>
- Misra, S. y Li, H. (2020). Noninvasive fracture characterization based on the classification of sonic wave travel times. In *Machine Learning for Subsurface Characterization* (pp. 243–287). Elsevier. <https://doi.org/10.1016/B978-0-12-817736-5.00009-0>
- Orlikowsld W. y Baroudi J. (1991). Studying Information Technology in Organimtions: Research Approaches and Assumptions. *Information Systems Research*.
- Osisanwo, F. Y. y Akinsola, J. E. T. y Awodele, O. y Hinmikaiye, J. O. y Olakanmi, O. y Akinjobi, J. (2017). Supervised Machine Learning Algorithms: Classification and Comparison. *International Journal of Computer Trends and Technology*, 48(3), 128–138. <https://doi.org/10.14445/22312803/IJCTT-V48P126>
- Padilla-Cuevas, J. y Reyes-Ortiz, J. A. y Bravo, M. (2021). Ontology-Based Context Event Representation, Reasoning, and Enhancing in Academic Environments. *Future Internet*, 13(6), 151. <https://doi.org/10.3390/fi13060151>
- Park, Y. S. y Konge, L. y Artino, A. R. (2020). The Positivism Paradigm of Research. *Academic Medicine*, 95(5), 690–694. <https://doi.org/10.1097/ACM.0000000000003093>
- Pastor, M. (1997). *Deep Blue venció a Kasparov en un campo donde el hombre no tenía rival* | *ComputerWorld*. Computer World. <https://www.computerworld.es/archive/deep-blue-vencio-a-kasparov-en-un-campo-donde-el-hombre-no-tenia-rival>
- Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 19–27.
- Salatino, A. A. y Thanapalasingam, T. y Mannocci, A. y Birukou, A. y Osborne, F. y Motta, E. (2020). The Computer Science Ontology: A Comprehensive Automatically-Generated Taxonomy of Research Areas. *Data Intelligence*, 2(3), 379–416. [https://doi.org/10.1162/dint\\_a\\_00055](https://doi.org/10.1162/dint_a_00055)
- Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, 2(3), 160. <https://doi.org/10.1007/s42979-021-00592-x>

- Saunders M. y Lewis, P. y Thornhill, A. (2019). *Research Methods for Business Students* (Pearson Education, Ed.; Eighth).
- Saunders, M. N. K. y Lewis, P. y Thornhill, A. (2029). *Research Methods for Business Students* (8va ed.).
- Scriven, Mi. (1967). *The methodology of evaluation*.
- Universidad de Lehigh. (2013). *Historia de las Pruebas Estandarizadas*.  
<https://ed.lehigh.edu/news-events/news/history-standardized-testing>
- Yeturu, K. (2020). *Machine learning algorithms, applications, and practices in data science* (pp. 81–206). <https://doi.org/10.1016/bs.host.2020.01.002>
- Žukauskas, P. y Vveinhardt, J. y Andriukaitienė, R. (2018). Philosophy and Paradigm of Scientific Research. In *Management Culture and Corporate Social Responsibility*. InTech. <https://doi.org/10.5772/intechopen.70628>

# CAPÍTULO 3

## Marco teórico conceptual

### Theoretical and conceptual framework

DOI: <https://doi.org/10.71112/rsbwvb15>

#### Resumen

Este capítulo explora la aplicación de algoritmos de aprendizaje automático —como Random Forest, Naive Bayes y Árboles de Decisión— para predecir el rendimiento académico. Se revisan antecedentes internacionales y nacionales que demuestran la efectividad de estas técnicas en la identificación temprana de estudiantes en riesgo. Además, se fundamenta en teorías educativas como la Taxonomía de Bloom y el Modelo de Productividad de Walberg, que destacan factores cognitivos, afectivos y ambientales clave en el aprendizaje. El marco integra conceptos de minería de datos educativos para mejorar la toma de decisiones en entornos académicos.

**Palabras clave:** Aprendizaje Automático, Rendimiento Académico, Minería de Datos Educativos, Taxonomía de Bloom, Modelo de Walberg, Predicción.

#### Abstract

This chapter explores the application of machine learning algorithms—such as Random Forest, Naive Bayes, and Decision Trees—to predict academic performance. It reviews international and national background studies demonstrating the effectiveness of these techniques in the early identification of at-risk students. Furthermore, it is grounded in educational theories like Bloom's Taxonomy and Walberg's Educational Productivity Model, which highlight key cognitive, affective, and environmental factors in learning. The framework integrates concepts from educational data mining to enhance decision-making in academic settings.

**Keywords:** Machine Learning, Academic Performance, Educational Data Mining, Bloom's Taxonomy, Walberg's Model, Prediction.

## **Antecedentes del problema**

### **a) Antecedentes internacionales**

Parhizkar et al., (2023), En la investigación realizada “Predicción del desempeño de los estudiantes utilizando algoritmos de clasificación de minería de datos: evaluación de la generalización de modelos desde el aspecto geográfico” Irán, describe que el incremento de la productividad en los sistemas educativos es de gran importancia, hoy en día los investigadores están interesados en predecir el rendimiento académico de los estudiantes; Esto se hace para mejorar la productividad general del sistema educativo mediante la identificación efectiva de los estudiantes cuyo rendimiento está por debajo del promedio. Esta preocupación universal se ha combinado con la ciencia de datos, lo que ha llevado a la creación de un área de investigación interdisciplinaria llamada Minería de Datos Educativos. Uno de los temas recientes que ha sido abordado por los investigadores es la formación de modelos generalizables desde diferentes aspectos como el género, la especialidad, la geografía, etc. Por lo tanto, en esta investigación utilizamos métodos de aprendizaje automático para predecir el rendimiento de los estudiantes, haciendo énfasis en el entrenamiento de modelos generalizables desde el punto de vista geográfico. Para ello, se diseñó un cuestionario de 37 preguntas, a través del cual se recogieron 536 respuestas, de las cuales 111 fueron internacionales y 425 nacionales. De acuerdo con la literatura, el rendimiento de los estudiantes se determina principalmente en función del GPA (promedio de calificaciones) de todo el curso. En esta investigación, se recopiló información sobre el promedio de los encuestados en cursos de pregrado y posgrado en forma de tres clases. Después de una revisión final de los modelos empleados en estudios previos, los principales modelos seleccionados y utilizados para fines de clasificación incluyeron SVM, CNN, Adaboost, RF, SVM y DT. La selección de características se realiza mediante XGBoost, bosque aleatorio y SVML1. La principal cuestión investigada en este estudio es la generalización de los modelos entrenados con datos nacionales (iraníes) y probados con datos internacionales (no iraníes). Los resultados experimentales muestran que los mejores modelos entrenados con conjuntos de datos específicos recopilados en esta investigación tuvieron generalización en comparación con los resultados de los modelos base que se entrenaron y probaron con datos nacionales. Por su parte, los modelos Random forest y CNN muestran el mejor rendimiento con un promedio de precisión y una puntuación F de 73,5 y 68,5, respectivamente.

Morales Hernández et al. (2022), realizaron un estudio titulado "Algoritmos de aprendizaje automático para la predicción del logro académico" en colegios ubicados en el estado de

Tlaxcala, México. En este estudio, se desarrollaron dos aplicaciones inteligentes: una red neuronal multicapa (perceptrón multicapa [MLP]) y un modelo de potenciación del gradiente (GB), con el objetivo de predecir el nivel de productividad académica en los cursos de español y matemáticas para alumnos del grado sexto del nivel primario y tercero de secundaria, cuyas variables pertenecieron a las evaluaciones nacionales del logro académico de los colegios escolares del estado mencionado de México. Es importante mencionar que se utilizaron 13 variables de entrada, cuya importancia fue determinada por el algoritmo de bosque aleatorio (RF). Las ententes inteligentes MLP y GB fueron entrenados y probados con un conjunto de datos de 11,036 registros pertenecientes a los escolares en el sistema escolar entre 2008 y 2011. Los modelos inteligentes ejecutaron predicciones en función de los registros de los periodos 2008 y 2011. En la asignatura de español, el modelo MLP superó al modelo GB con una precisión global de clasificación (PG) de 70.1 % en 2008 y 61.1 % en 2011. En referencia al curso de matemáticas, el modelo GB mostró mejores resultados, con una PG de 68.8 % en 2008 y 63.5 % en 2011. Asimismo, se pudo identificar una fuerte correlación entre el puntaje en del curso español y el grado de logro académico en el curso de matemáticas. Es importante destacar que los puntajes en español y matemáticas fueron relevantes, en comparación con factores contextuales como el sexo, beca y el turno escolar. Conforme a la muestra analizada, se observó que, en ambas asignaturas, el sexo femenino superó al sexo masculino en los grados de logro académico elemental, bueno o excelente; pero se invertiría en el grado de logro insuficiente.

Yağcı (2022), en su modelo de estudio denominado Minería de datos educativos: predicción de la productividad académica de los estudiantes de una universidad estatal turca mediante algoritmos inteligentes, propone predecir calificaciones de los exámenes finales de los estudiantes de pregrado, en donde referencia las calificaciones de las evaluaciones parciales para dicho estudio, asimismo utiliza algoritmos ML Random Forest, Vecinos más Cercanos, Regresión Logística, Naive Bayes y algoritmos de k-vecino más cercano para dicha investigación, para ello utilizó una muestra de 1,854 estudiantes del curso de idioma turco I en una universidad estatal en Turquía, durante el semestre de otoño de 2019-2020, cuyo resultado del modelo propuesto logró una precisión de clasificación del 70% al 75%; cabe mencionar que para las predicciones consideró tres tipos de parámetros calificaciones de las evaluaciones parciales, datos del departamento y datos de la facultad, en donde atribuye que dichos datos son muy importantes para el análisis de aprendizaje en la educación superior a fin de contribuir los procesos de toma de decisiones. El autor menciona que la contribución de la investigación está orientada a la predicción temprana de estudiantes con alto grado de fracaso y determina los métodos de aprendizaje automático más efectivo.

Nedeva y Pehlivanova (2021), realiza el trabajo de investigación “Análisis de rendimiento de los estudiantes utilizando algoritmos de aprendizaje automático en WEKA”, cuyo objetivo es pronosticar la productividad de los estudiantes de la Universidad Trakia de Stara Zagora de Bulgaria, con la finalidad de apoyar y mejorar los resultados educativos, asimismo identifica características más significativas (12 atributos con impacto en el desempeño de los estudiantes) que afectan el desempeño de los estudiantes, luego eligen el algoritmo inteligente más eficiente para predecir su desempeño. Se comparó la eficiencia de cuatro algoritmos de clasificación: BayesNet (BN), Perceptrón multicapa (MLP), Optimización mínima secuencial (SMO) y Árbol de decisión (J48), los mejores resultados se obtuvieron utilizando el algoritmo Multilayer Perceptron MLP, aplicado a 12 atributos, con valores principales: Tasa TP = 0,974; Precisión = 0,976, Medida F = 0,974 y Exactitud = 97,39 %, el cual concluyo con una mejor precisión en el pronóstico.

Yildiz y Börekci (2020), en la investigación realizada acerca de la Predicción de la productividad académica de los estudiantes mediante técnicas de aprendizaje automático, desarrolla una visión de los datos educativos recopilados en estudiantes del noveno grado, para ello se utilizaron datos con información demográfica de estudiantes y familias, rutinas de estudio, comportamientos de asistencia y creencia epistemologías sobre la ciencia, dicha investigación tuvo como objetivo resolver un problema de clasificación para ello intenta estimar clases de aprobado y desaprobado en los exámenes, asimismo dicho estudio hizo una predicción comparativa de precisión con los algoritmos de clasificación supervisada, definiendo que variables eran más efectivas en la formación académica, adicionalmente se examinó el coeficiente de ganancia de información de las variables para determinar los factores que afectan la precisión de la predicción.

En el estudio, se comparó la precisión de predicción de los algoritmos de clasificación supervisada y se definió qué variables eran efectivas en la formación de clases. Cuando se comparó la precisión de predicción de los algoritmos inteligentes, los resultados indicaron que el algoritmo de red neuronal (98,6 %) obtuvo la puntuación más alta. La tasa de precisión de los otros algoritmos son K vecinos más cercanos (KNN) (86,2 %), regresión logística (78,4 %), SVM (90,3 %), árbol de decisión (91,9 %), bosque aleatorio (Random forest) (90,0 %) e Naive Bayes (81,7 %). Se examinó el coeficiente de ganancia de información de las variables para determinar los factores que afectan la precisión de la predicción. Se puede concluir que existió una relación entre estas variables y el éxito académico. Los estudios sobre estas variables apoyarán el éxito académico de los estudiantes. Finalmente, el autor recalca que las variables de investigación apoyaran el éxito académico de los estudiantes.

## **b) Antecedentes nacionales**

Gismondi (2021), en la investigación titulada “*Modelo predictivo basado en machine learning como soporte para el seguimiento académico del estudiante universitario*”, cuya finalidad fue mejorar los resultados en la educación universitaria, propone aplicar la inteligencia artificial, machine learning y Deep learning, para ello considera atributos relevantes a fin de proponer un modelo de predicción de aprendizaje profundo con el algoritmo de redes neuronales con 2, 3, 4, 5, 6 y 7 capas con el respectivo entrenamiento y prueba, para ello se concluyó que se obtuvo mejores resultados con redes neuronales de 6 capas con un 98.97% de precisión en el entrenamiento y 81.73 de precisión en el conjunto de prueba.

Gismondi y Huiman (2021), elaboran la investigación “*Características para un modelo de predicción de la deserción académica universitaria. Caso Universidad Nacional del Santa*”, proponen un *modelo* predictivo a fin de disminuir la deserción académica de estudiantes de la Universidad Nacional de Santa, para ello propone atributos que formaran parte de un modelo inteligente basado en Machine Learning, para ello se realizó una revisión documental y métodos estadísticos; finalmente concluyen la propuesta con seis características académicas y doce indicadores dentro de las características no académicas.

Díaz-Landa et al. (2021), propone la investigación realizada “*Rendimiento académico de estudiantes en Educación Superior: predicciones de factores influyentes a partir de árboles de decisión*”, cuyo objetivo es predecir el rendimiento académico de estudiantes de maestrías en educación, para ello empleó técnicas de árbol de decisión, minería de *datos* y algoritmo J48 en la herramienta WEKA, asimismo considero dimensiones educativas, familiares, socio económicos, hábitos personales y costumbres de una población de 237 estudiantes de una universidad pública Peruana, luego efectuó el entrenamiento del modelo, logrando establecer que el algoritmo de clasificación bayesiano J48 una efectividad del 66.6%, identificando asimismo mediante técnicas inteligentes los factores influyentes en el rendimiento académico: enseñanza, convenientes horarios de clase, adecuada relación social docente-estudiante, nivel académico.

Quiñones y Quiñones (2020), llevaron a cabo una investigación titulada “*Rendimiento académico empleando minería de datos*”, cuyo propósito fue predecir el rendimiento académico de los estudiantes de la especialidad de Industrias Alimentarias en la Universidad Nacional de Jaén, mediante modelos de minería de datos. Los datos históricos en el estudio fueron recopilados a través de formularios y registros de las oficinas dicha universidad. La metodología aplicada fue CRISP-DM; cuyos pronósticos de los modelos inteligentes se realizaron en el software weka cuyos resultados fueron superiores al 83% de efectividad.

Chiok (2017), presenta una investigación titulada "Predicción del rendimiento académico aplicando técnicas de minería de datos, curso de Estadística General de la Universidad Nacional Agraria La Molina (UNALM)", donde acentúa que el rendimiento académico de los estudiantes es una de las principales preocupaciones para las instituciones de educación superior. El estudio destaca que las técnicas de minería de datos (TMD) aplicadas a entornos educativos se han convertido en una herramienta eficaz para pronosticar la productividad académica, a fin de identificar los factores más influyentes en el aprendizaje de los estudiantes y apoyar a los docentes a optimar el proceso enseñanza aprendizaje con acciones más eficaces y oportunas. La investigación aplico TMD como el algoritmo de regresión logística, algoritmo árboles de decisión, algoritmo de redes bayesianas y el algoritmo de redes neuronales, cuyos datos de entrenamiento pertenecieron a los estudiantes de la asignatura de estadística general de UNALM durante los semestres académicos 2013-II y 2014-I, con la finalidad de pronosticar la evaluación final de aprobación o reprobación del futuro en estudiantes del curso. Los resultados del estudio revelan que el algoritmo Naive Bayes alcanzó la mayor tasa de efectividad, con un 71.0% de precisión.

## **Bases teóricas o científicas**

### **a) Teoría de la productividad educativa**

#### **Taxonomía de bloom de los objetivos de aprendizaje**

Según Gogus (2012), la taxonomía es el proceso científico de clasificar y organizar; asimismo en el sentido académico los objetivos de aprendizaje es el logro y demostración que se espera del estudiante luego de un periodo de tiempo. La taxonomía de Bloom desarrolla una jerarquía de objetivos educativos que **contenía categorías en el dominio cognitivo**, en donde menciona al conocimiento, luego a la comprensión, seguido de la aplicación, posteriormente al análisis, seguido de la síntesis y finalmente a la evaluación (Amer, 2006, como se citó en Rosales Sánchez et al., 2020).

Sepúlveda et al. (2020) menciona que la taxonomía de Bloom se convirtió en un componente esencial para la estructura y comprensión de las actividades educativas basadas en el proceso enseñanza aprendizaje en donde identifiqué dominios de las actividades educativas que a continuación se describen:

- **Dominio cognitivo:** Adquisición de conocimientos y habilidades intelectuales (Conocimiento).
- **Dominio afectivo:** Se refiere a la integración de credos e ideas, fundamentada en las actitudes y sentimientos que participan en el proceso. (Actitud).



- **Dominio psicomotor:** Basado en la adquisición de movimientos coordinados, tanto manuales como físicos (Habilidades).

Bloom también define dominios y subdominios dentro de las seis categorías principales (niveles), organizándolas de lo simple a lo complejo y a partir de lo definido hacia lo abstracto (Krathwohl, 2002). Además, se asumía que el dominio de una categoría más sencilla era esencial para poder dominar la siguiente categoría de mayor complejidad (Ver Tabla 1).

**Tabla 1**

*Taxonomía de Bloom de los objetivos de aprendizaje para el dominio cognitivo*

Dominio	Niveles
Cognitivo simple o bajo	1. <b>Conocimiento</b> , capacidad de memorizar y recordar hechos específicos.
	2. <b>Comprensión</b> , capacidad de interpretar y demostrar comprensión básica.
Cognitivo superior	3. <b>Aplicación</b> , capacidad de aplicar conceptos y separar conceptos en componentes.
	4. <b>Análisis</b> , capacidad de analizar conceptos
Habilidades del pensamiento crítico.	5. <b>Síntesis</b> , capacidad para combinar elementos y formar un todo.
	6. <b>Evaluación</b> , capacidad de emitir juicios sobre el valor de lo aprendido.

Nota: extraído de (Gogus, 2012)

Asimismo, Serumena D. et al. (2021), destaca que el dominio afectivo está caracterizado por las emociones, valores, apreciación, entusiasmo, interés, motivación y actitudes, dichas categorías están ordenadas del más simple al más complejo:

**1ro. Recepción.** capacidad de prestar atención y responder a estimulación adecuada, la aceptación es el más bajo.

**2do. Sensible,** los estudiantes se involucran de manera afectiva, participan y toman interés.

**3ro. Valor adherido,** se refiere al valor o importancia de atracción a ciertos objetos o eventos con reacciones a la aceptación, rechazo o ignorar el objeto. El objetivo está basado en “actitud / aprecio”, en la valoración de lo bueno y lo malo.

**4to. Organización**, centrado a la unificación de valores, actitudes que generan conflictos internos que va determinando el comportamiento que es representado en una filosofía de vida y la capacidad de formar un sistema de valores y cultura organizacional.

**5to. Caracterización**, se refiere al carácter y la fuerza vital de la persona, considerando los valores están evolucionados para que el comportamiento se vuelva consistente y predecible.

Finalmente, los dominios psicomotores incluyen el aspecto físico y coordinación, habilidades motoras y capacidades físicas:

**1ro. Imitación**, cuando empieza a responder de manera similar a los observados.

**2do. Manipulación**, se refiere a la habilidad de seguir direcciones, apariencias, movimientos seleccionados determinando una apariencia a través de la práctica.

**3ro. Idoneidad**, adquiere mayor precisión, proporción y certeza en la apariencia.

**4to. Articulación**, establece coordinación de movimientos de manera exacta y logrando consistencia esperada entre diferentes movimientos.

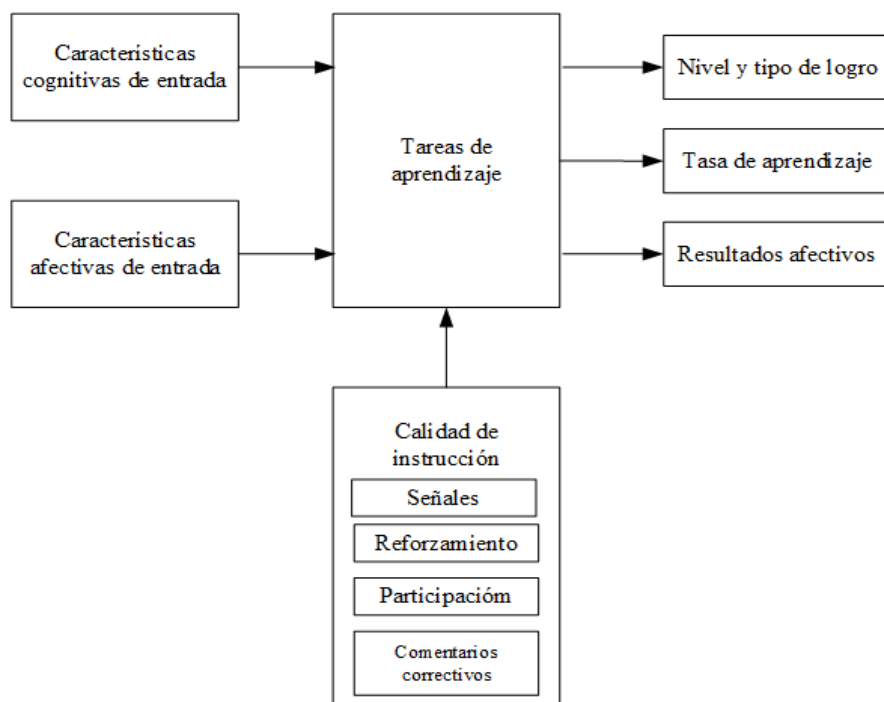
**5to. Experiencia**, su conducta expresa menor gasto energético físico y psíquico.

### **Modelo de aprendizaje bloom**

Según, Seel (2012) el modelo de aprendizaje de Bloom propuesto en 1976 , tiene una característica de carácter cíclico de la instrucción, lo cual significa que los resultados obtenidos de una tarea, se convierte en el aporte del anterior; el modelo está representado por cuatro dimensiones: características cognitivas de entrada, características afectivas de entrada, calidad de instrucción, los resultados de aprendizaje se especifican por el nivel y tipo de logro, tasa de aprendizaje o resultados afectivo (Figura 1).

### **Figura 1**

*Modelo de aprendizaje de Bloom*



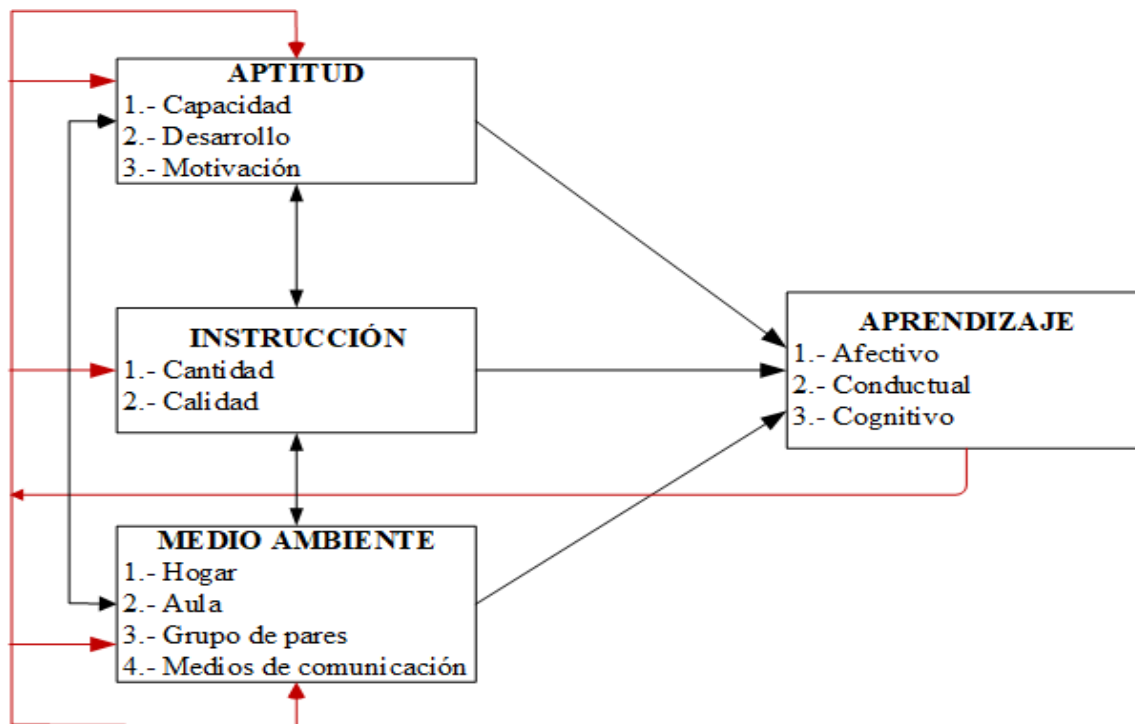
Nota: Extraído de (Seel, 2012a)

### Modelo de la productividad educativa

Modelo propuesto por Walberg (1984), basado en el rendimiento académico del estudiante, propone cuatro componentes (Figura 2) que se combinan e influyen entre sí, con la finalidad de determinar el rendimiento final del estudiante, *asimismo recalca si el estudiante tiene habilidades favorables y ambientes favorables bajo estímulos, entonces el aprendizaje puede ser positivo en términos de actitudes, comportamientos y conocimientos*. Es importante considerar que la teoría de Walberg menciona que las características psicológicas influyen en los resultados de aprendizaje individual del estudiante en lo cognitivo, conductual y actitudinal (Mazana et al., 2019), (Peconcillo Jr et al., 2020), por tanto dichas características psicológicas influyen directamente en el rendimiento académico (Reynolds y Walberg, 1992; Wang et al., 1993, como se citó en Abid et al., 2021), es decir que las nueve variables (grafico 4) contribuyen al rendimiento académico (Walberg y Tsai, 1985; Wang et al., 1993, como se citó en Abid et al., 2021).

**Figura 2**

*Modelo de productividad educativa*



*Nota:* obtenido de (Walberg, 1984)

### **a) Aptitud del estudiante**

La aptitud se refiere a las habilidades y capacidades innatas de los estudiantes. Incluye tanto las habilidades cognitivas (como la inteligencia, el razonamiento lógico y la memoria) como las habilidades no cognitivas (como la motivación, la perseverancia y la autoestima). El comportamiento del estudiante puede influir en su aptitud al afectar su nivel de compromiso, su actitud hacia el aprendizaje y su disposición para superar desafíos, asimismo, implica:

- **Capacidad:** Se refiere a la habilidad innata de un individuo para aprender y comprender nuevos conceptos. Los indicadores de capacidad pueden incluir la rapidez con la que se adquieren nuevas habilidades cognitivas.
- **Motivación:** La disposición y el interés del estudiante para participar activamente en el proceso de aprendizaje. Los indicadores de motivación podrían incluir la persistencia en la resolución de problemas y el entusiasmo por aprender.

- **Desarrollo:** Se refiere al progreso y crecimiento del estudiante a lo largo del tiempo. Indicadores de desarrollo pueden ser la mejora continua en el desempeño académico y el desarrollo de habilidades sociales.

#### **b) Instrucción**

Se refiere al tiempo que el estudiante se involucra en el aprendizaje. Asimismo, a la calidad de experiencia educativa, que incluye aspectos psicológicos y curriculares, los cuales involucra:

- **Cantidad:** La cantidad de tiempo y recursos dedicados a la enseñanza. Indicadores de cantidad pueden incluir la duración de las clases, la cantidad de material didáctico proporcionado y el tiempo total de instrucción.
- **Calidad:** La efectividad de la enseñanza. Indicadores de calidad podrían abarcar la claridad en la presentación de conceptos, la retroalimentación constructiva y la adaptación de métodos de enseñanza a las necesidades individuales de los estudiantes.

#### **c) Medio Ambiente**

Son factores constantes que afectan al aprendizaje como el clima psicológico y estímulos educativos: el hogar, grupo social del aula, compañeros fuera de la institución educativa, usos del tiempo extra escolar (actividades libres de realizar) los cuales involucra:

- **Hogar:** El entorno familiar en el que crece el estudiante. Indicadores del hogar podrían incluir el apoyo de los padres, la disponibilidad de recursos educativos en casa y la estabilidad del entorno familiar.
- **Entorno comunitario:** La influencia del entorno fuera del hogar. Indicadores del entorno comunitario pueden abarcar el acceso a servicios educativos adicionales, la presencia de modelos a seguir en la comunidad y la seguridad del entorno.
- **Medios de comunicación:** El impacto de los medios de comunicación en la educación. Indicadores de medios de comunicación podrían incluir el acceso a recursos educativos a través de medios electrónicos, la calidad de los programas educativos y el tiempo dedicado a actividades educativas frente a las pantallas.

#### **d) Aprendizaje**

En esta etapa está referido a los logros obtenidos o conocimientos nuevos durante el proceso del aprendizaje como un proceso de producción y pueden ser de distintos tipos afectivo, conductual o cognitivo. El aprendizaje será mayor cuando el ambiente es cooperativo con

orientaciones claras y con objetivos y con estímulos necesarios para el aprendizaje, cuyo indicador resaltante es:

- **Logro:** de detalla el valor del nivel de éxito alcanzado por el estudiante en términos de conocimientos adquiridos y habilidades desarrolladas. Indicadores de logro pueden abarcar calificaciones académicas, resultados en pruebas estandarizadas y la aplicación exitosa de conocimientos en situaciones prácticas.

### **Modelo de aprendizaje escolar de carroll**

El modelo de Carroll propuesto en 1960 (Figura 3) destaca los roles distintivos de las habilidades, aptitudes para el logro del aprendizaje; la eficacia se determina en función del tiempo para aprender y del tiempo realmente dedicado al aprendizaje (Guill et al., 2021; Rosales Sánchez et al., 2020).

Ecuación funcional de grado de aprendizaje:

$$\text{Grado de aprendizaje} = \frac{\text{tiempo real necesario para el aprendizaje}}{\text{tiempo realmente necesario para el aprendizaje}}$$

Para ello referencia al aprendizaje como activo cuantificado del éxito logrado durante un periodo fijo de tiempo, el mismo que puede variar dicha tarea de aprendizaje desde minutos, días, semanas o meses para tareas complejas, asimismo, dependerá de otras variables: comprensión y aptitud del estudiante para el logro del encargo académico (Seel, 2012b).

Asimismo, Seel, menciona que Carroll (1989) establece la existencia de cinco variables en dos clases principales concebidas para influir en el grado de aprendizaje; ello se resume de la siguiente manera:

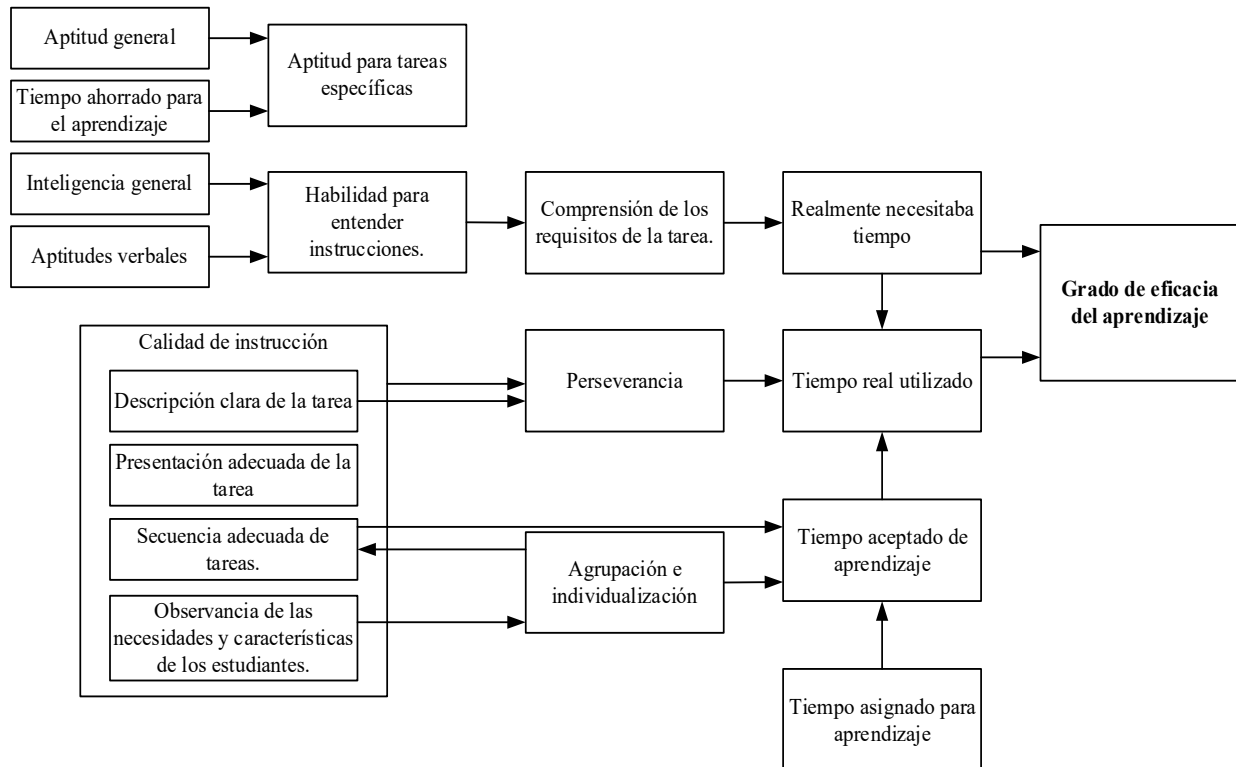
Variables instructivas:

- Calidad de la instrucción (trabajo presentado)
- Oportunidad (tiempo de aprendizaje)

Variables de diferencias individuales

- Aptitud del estudiante para aprender la tarea en función del tiempo, capacidad y logros previos.
- Capacidad de entendimiento de la instrucción (inteligencia y habilidad verbal)
- La "perseverancia" del estudiante: el tiempo como inversión de aprendizaje.

### Modelo de aprendizaje Carroll



*Nota:* obtenido de (Seel, 2012b)

## b) Teoría de algoritmo de aprendizaje

La inteligencia es la capacidad de aprender, entender y pensar en las cosas. Es una capacidad intelectual general que implica la capacidad de razonar, planificar, resolver complicaciones, pensar abstractamente, percibir ideas y aprender (Sauce et al., 2022). Asimismo, tiene la capacidad para llegar a la salida más adecuada a pesar de todas las probabilidades en una variedad de situaciones (Vostroknutov et al., 2018); puede adquirir conocimientos para superar dificultades y producir resultados apropiados presentados (Zhao, 2022); por otro lado, la Inteligencia Artificial (IA) se puede determinar cómo inteligencia ejecutada por un sistema fabricado o hecho por el hombre. Es un intento de hacer que una máquina se comporte de maneras que se llamarían inteligentes si un humano hiciera lo mismo. La Inteligencia Artificial (IA) es un campo diferente donde los investigadores persiguen una amplia gama de complicaciones, utilizan una selección de métodos y persiguen un espectro de objetivos

científicos, según Wang (2019) se debe establecer cuatro criterios para una definición funcional de la IA basado en especificar su uso común, establecer un límite claro, conducir a una investigación exitosa y debe ser lo más simple posible el desarrollo un sistema inteligente. El objetivo principal de un sistema basado en Inteligencia Artificial (IA) es alcanzar un nivel humano de inteligencia(Vigo et al., 2022), (Dong et al., 2020); a pesar de existir diferentes definiciones, el entendimiento común de la IA está enfocada a máquinas y computadoras con la finalidad de ayudar a la humanidad a resolver problemas y facilitar los procesos de trabajo (Tai, 2020).

En la rama de IA encontramos el enfoque del aprendizaje automático (Machine Learning (ML)) basado en una colección de algoritmos que intentan extraer patrones de los datos y asociar dichos patrones con clases discretas de muestras en los datos, asimismo dentro del enfoque encontramos métodos como ML supervisado, no supervisado y por refuerzo, que intentan encontrar una secuencia de acciones con la finalidad de lograr un objetivo específico (Jovel y Greiner, 2021), Para (Kelleher, Mac Namee, & D'Arcy, 2015), el aprendizaje automático es una disciplina científica que capacita a las máquinas para desarrollar técnicas de aprendizaje. Este proceso automatizado extrae patrones de los datos y construye modelos que posibilitan la predicción mediante algoritmos supervisados, estableciendo conexiones entre características descriptivas actuales y objetivos basados en un conjunto de ejemplos o instancias históricas. Este proceso consta de dos pasos principales: data set, aprendizaje supervisado y su posterior pronóstico.

Por otra parte (Mitchell, 1997), establece que el aprendizaje automático se fundamenta en conceptos de diversas disciplinas, como la disciplina de la inteligencia artificial, la disciplina de la probabilidad y la estadística, la disciplina de la complejidad computacional, la teoría de la información, teoría de psicología y neurobiología, teoría del control y la disciplina de la filosofía. Según Jovel y Greiner (2021), la implementación de estos conceptos a través de algoritmos computacionales mejora el rendimiento en función de la experiencia de aprendizaje; para Janiesch et al. (2021) los algoritmos inteligentes tienen la capacidad para aprender de los datos conforme a entrenamientos específicos del problema a fin de automatizar el proceso de creación del modelo y posteriormente resolver tareas asignadas.

Según Bangato J. I. (2020) menciona que es una subdisciplina de la Inteligencia Artificial basado en “cómo construir programas de computadora que mejoran automáticamente adquiriendo experiencia”, menciona que la implementación desarrollada con ML no necesita de reglas o funciones explícitas del lenguaje de programación, sino que se desarrolla de manera automática, asimismo, destaca de los algoritmo inteligentes utilizan pocos recursos durante el



proceso de entrenamiento de grandes volúmenes de datos y cuyo aprendizaje se destaca de manera autónoma. Clasifica al aprendizaje automático en dos tipos: una referida al aprendizaje supervisado o al aprendizaje no supervisado y la otra refiere al aprendizaje por reforzamiento de los algoritmos.

Podemos mencionar que los algoritmos inteligentes más utilizados son: los algoritmos de árboles de decisión, algoritmos de regresión lineal, algoritmos de regresión logística, algoritmos k Nearest Neighbor, algoritmos PCA / Principal Component Analysis, algoritmos SVM, algoritmos gaussian Naive Bayes, algoritmos K-Means, algoritmos redes neuronales artificiales, algoritmos de aprendizaje profundo.

### **Aprendizaje Supervisado**

Según, Jiang et al. (2020) el aprendizaje supervisado consiste en métodos diseñados para predecir o clasificar un resultado acorde a un objetivo específico, asimismo el aprendizaje se utiliza para describir tareas de predicción porque el objetivo es pronosticar/clasificar un resultado específico de interés; por otro lado, Stern et al.(2020) sostiene que los algoritmos de aprendizaje supervisado pueden aprender características sutiles que distinguen una clase de ejemplos de entrada de otra; adicionalmente, Pugliese et al. (2021) menciona que el aprendizaje supervisado se basa en tareas para aprender una función establecida, la misma que ha sido asignada una entrada a una salida en función de pares de entrada-salida de la muestra; dicho proceso de aprendizaje está basado en datos representados por un dominio de datos con similitudes genéricas (Münch et al., 2020).

Según Lampropoulos y Tsihrintzis (2015), el aprendizaje supervisado se encarga de tareas predictivas, de clasificación y constituye la categoría principal de aprendizaje. En el aprendizaje supervisado, los objetos que están relacionados con un concepto específico son pares de patrones de entrada y salida. Esto significa que los datos que pertenecen al mismo concepto ya están asociados con valores objetivo, como, por ejemplo, clases que definen las identidades de conceptos, por su parte Bagnato J.I. (2020), en el aprendizaje supervisado, los datos de entrenamiento contienen etiquetas, o la respuesta deseada. Determinar si el correo entrante es spam es un buen ejemplo. Una de las muchas características que deseamos entrenar debe ser si la entrada es spam, denotada por un 1 o un 0. Otro ejemplo es cuando tenemos que añadir el precio que descubrimos en nuestro conjunto de datos y tenemos que estimar valores numéricos, como el precio de una casa, basándonos en sus características (metros cuadrados, número de habitaciones, calefacción, distancia al centro, etc.).

El algoritmo k-vecinos más cercanos, algoritmo regresión lineal, algoritmo regresión logística, algoritmo máquinas de vectores de soporte, algoritmo clasificadores bayesianos, algoritmo arboles de decisión, algoritmo bosque aleatorio, algoritmo redes neuronales y los algoritmos de aprendizaje profundo son los algoritmos más utilizados en el aprendizaje supervisado.

### **Aprendizaje No Supervisado**

El modelo no supervisado según Lee et al. (2022) es el sistema que desarrolla el aprendizaje sin intervención humana para predecir; dicho modelo no construye ningún modelo predictivo, para la evaluación de nuevos datos, existen dos opciones; ya sea el mapeo en el espacio agrupado o el agrupamiento o la reducción de la dimensión se realiza con todos los datos una vez más (Tebani y Bekri, 2020); los algoritmos no supervisados se utilizan con frecuencia para determinar las etiquetas de los grupos de datos que no tienen etiquetas (KOÇOĞLU, 2022).

Según Zheng (2019) el desarrollo del aprendizaje no supervisado se basa en agrupar los datos conglomerados donde dichos elementos son muy similares y los elementos de diferentes conglomerados muestran diferencias mucho mayores. Debido a que los elementos no tienen etiquetas de clase, que se conocen en tareas de clasificación o aprendizaje supervisado, la agrupación en clústeres se reconoce como aprendizaje automático no supervisado.

Asimismo, Lampropoulos y Tsihrintzis (2015) explica que se trata de comprender o encontrar una descripción concisa de datos mediante el mapeo pasivo o la agrupación de datos de acuerdo con algunos principios de orden. Esto significa que los datos constituyen solo un conjunto de objetos donde no se dispone de una etiqueta para definir el concepto asociado específico como en el aprendizaje supervisado. Así, el objetivo del aprendizaje no supervisado es crear grupos-clusters de objetos similares de acuerdo con un criterio de similitud y luego inferir un concepto que se comparte entre estos objetos. Asimismo, el aprendizaje no supervisado incluye algoritmos que tienen como objetivo proporcionar una representación de espacios de alta a baja dimensión, preservando la información inicial de los datos y ofreciendo un cálculo más eficiente. Por su parte (Bagnato, 2020) establece que a diferencia del aprendizaje supervisado, el aprendizaje no supervisado no incluye etiquetas predeterminadas con los datos de entrenamiento. Más bien, el algoritmo se encarga de ordenar y evaluar automáticamente los datos, encontrando estructuras o patrones ocultos en la información. Este método es utilizado comúnmente en la segmentación de usuarios en sitios web o aplicaciones, donde los usuarios son agrupados por la inteligencia artificial basándose en características compartidas que el sistema reconoce por sí mismo. La agrupación de K-Means, el análisis de componentes principales y la detección de anomalías son algunos de los algoritmos más

importantes del aprendizaje no supervisado; todos ellos son esenciales para identificar patrones subyacentes y descifrar datos complejos sin necesidad de supervisión humana.

### **Aprendizaje por refuerzo**

Según, Estes et al. (2022) el aprendizaje por refuerzo es similar a la forma en que aprenden los humanos y los animales, es decir que el aprendizaje por refuerzo dentro del contexto más amplio de la inteligencia artificial permite que los agentes algorítmicos reemplacen a los seres humanos en el mundo real, destacando su participación en diferentes escenarios casas, edificios etc., cuyos dominios hasta ahora consideradas más allá de las capacidades actuales (Al-Ani y Das, 2022); por otro lado Xiang et al. (2021) menciona que el aprendizaje por refuerzo (RL) es un enfoque para simular el proceso de aprendizaje natural del ser humano, es decir imita el proceso de aprendizaje de los humanos, entrena cometiendo y luego evitando errores,, puede resolver algunos problemas que los métodos convencionales no pueden resolver, en algunas tareas, también tiene la capacidad de superar a los humanos; un aspecto estable según, Singh et al.(2022) es que el sistema de aprendizaje por refuerzo, no se proporcionan pares de entrada y salida, al estado actual del sistema se le asigna una meta específica y un conjunto de acciones permitidas y restricciones para esperar sus resultados.

Según, Lampropoulos y Tsihrintzis (2015), el aprendizaje por refuerzo implica realizar acciones para lograr un objetivo, el agente aprende por ensayo y error al realizar una acción, asimismo espera una recompensa; de esta manera el método es eficiente para desarrollar estrategias de acción dirigidas a objetivos. El aprendizaje por refuerzo se inspiró en teorías psicológicas relacionadas y está fuertemente relacionado con los ganglios basales del cerebro. Las metodologías de aprendizaje por refuerzo se relacionan con problemas donde el agente de aprendizaje no sabe a priori lo que debe hacer, por lo tanto, el agente debe descubrir una política de acción que maximice la “ganancia esperada”. El aprendizaje por refuerzo difiere del aprendizaje supervisado porque en el aprendizaje por refuerzo no se presentan pares de entrada / salida, ni se corrigen explícitamente acciones subóptimas, sino que los agentes en un momento específico caen en un estado y con esta información seleccionan una acción, como consecuencia, el agente recibe por su acción una señal de refuerzo o recompensa; finalmente, Bagnato Juan Ignacio (2020), describe que durante este tipo de sistemas, el agente se comporta como un “agente autosuficiente” que debe descubrir un entorno desconocido, denominado “espacio”, para decidir los movimientos a realizar mediante un procedimiento de ensayo y error. A través de este procedimiento, el agente aprende de forma independiente, recibiendo recompensas por los movimientos correctos y consecuencias por los errores, lo que le permite

averiguar la forma más beneficiosa de llevar a cabo responsabilidades como recorrer un camino, resolver un rompecabezas o tal vez cultivar técnicas de recreación en entornos como Pac-Man o Flappy Bird. Este aprendizaje continuo permite al agente formular la mejor estrategia viable, denominada “política”, que le ayudará a tomar decisiones para maximizar las recompensas en el menor tiempo posible y de la forma más ecológica posible. Las políticas, por tanto, definen los movimientos que deben realizarse en cada situación concreta que se le presente al agente. Entre las máximas modas usadas habitualmente en este método está Q-Learning, que se basa totalmente en los Procesos de Decisión de Markov (MDP), esenciales para la toma de decisiones en entornos en los que no siempre se dispone de registros completos y el agente debe adaptarse continuamente a las circunstancias cambiantes.

### c) Pasos para construir un modelo de machine learning

Simplilearn (2022), señala que la creación de un modelo de aprendizaje automático es un proceso que suele constar de las siguientes fases y no se limita al uso de algoritmos o bibliotecas de aprendizaje:

- **Recolectar los datos.** La información puede obtenerse de diversas fuentes, como bases de datos, sitios web, API y otros recursos en línea. También podemos utilizar datos que sean de dominio público o utilizar otros dispositivos que recopilen los datos por nosotros. Las posibilidades que tenemos para recopilar datos son infinitas. Aunque este proceso parece sencillo, es uno de los más complicados y de los que más tiempo requieren.
- **Preprocesar los datos.** Tras obtener los datos, debemos asegurarnos que dichos datos sean homogéneos, estén limpios y tienen el formato correcto para poder alimentar nuestro algoritmo inteligente. Antes de poder utilizar los datos, es casi seguro que tendremos que completar una serie de operaciones de preparación. Pero esta parte suele ser mucho más fácil que la anterior.
- **Elegir el modelo.** Un modelo de aprendizaje automático determina el éxito del resultado que se obtiene luego de ejecutar el algoritmo con los datos recopilados. Es importante elegir un modelo relevante para desarrollar dicha tarea en cuestión. Recordemos que existen muchos aportes de modelos inteligentes significativos por parte de científicos e ingenieros para diferentes tareas como reconocimiento de voz, reconocimiento de imágenes, predicción, etc. Aparte de esto, también debe ver si su modelo es adecuado para datos numéricos o categóricos y elegir en consecuencia.

- **Entrenar el algoritmo.** La capacitación del entrenamiento, se basa en preparar los datos para el algoritmo inteligente a fin encontrar patrones y realizar pronósticos, dando como resultado que el modelo aprenda de los datos para que pueda realizar el conjunto de tareas. La idea es que, a medida que pase el tiempo, los algoritmos sean capaces de “aprender” información de los datos que proporcionamos, lo que permitirá al modelo hacer predicciones más precisas.
- **Evaluar el algoritmo.** En este punto, evaluamos el rendimiento del modelo utilizando un nuevo conjunto de datos que no se utilizó durante el entrenamiento. Esto es importante, ya que, si se utilizara el mismo conjunto de datos, el modelo se limitaría a repetir patrones aprendidos previamente, lo que daría lugar a una precisión engañosamente alta. Podemos determinar el rendimiento del modelo en escenarios reales y su capacidad de generalización con mayor precisión probándolo con datos que no se hayan visto antes.
- **Ajuste de parámetros.** Luego de la creación y evaluación del modelo, se debe verificar la precisión, esto se hace ajustando los parámetros en el modelo. Los parámetros son las variables en el modelo que el programador generalmente decide, es un valor particular que se debe ajustar a fin de obtener la precisión máxima.
- **Utilizar el modelo.** Por último, el modelo se aplica al problema real, lo que permite evaluar su eficacia en un entorno real. Es posible que haya que revisar y modificar los procesos anteriores a la luz de los resultados.

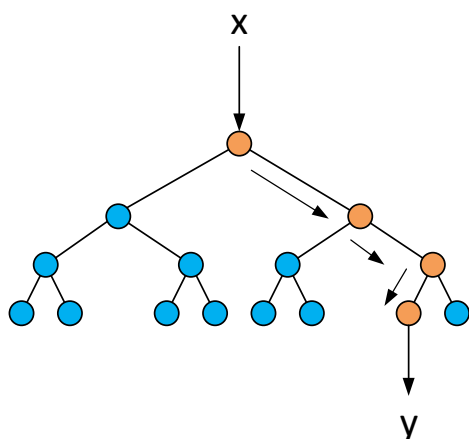
#### **d) Teoría de algoritmos de aprendizaje automático**

##### **Árboles de decisión**

Las tareas de regresión y clasificación son algunas de las muchas funciones que pueden ejecutar los árboles de decisión. Se trata de algoritmos extremadamente potentes que pueden adaptarse a conjuntos de datos modernos y complicados (Géron, 2019); según Hajjej et al. (2022) el objetivo principal del algoritmo es el aprendizaje inductivo basado en la observación y el razonamiento lógico, se utiliza el árbol para simbolizar la información del proceso de aprendizaje inductivo; los árboles se pueden representar mediante nudos, hojas y ramas en una representación visual, cuya clasificación comienza con un nodo raíz que representa el atributo del que se restan todos los demás atributos; asimismo (Mienye et al., 2019), describe que una vez construido el árbol de decisión se utilizan para clasificar los datos de prueba; manifiesta que el conjunto de datos diseñado para entrenar el árbol de decisión sirve como entrada para

Russell y Norvig (2008), una de las mejores técnicas para crear algoritmos de aprendizaje, según su libro, es el árbol de decisión. Este método utiliza como entrada un conjunto de atributos para caracterizar un objeto o circunstancia y, a continuación, emite una "decisión", es decir, un valor predicho basado en la entrada. Aunque en este contexto nos centraremos en las entradas discretas, las cualidades de entrada pueden ser continuas o discretas. El procedimiento se denomina clasificación cuando se aprende una función de valores discretos y regresión cuando se trabaja con una función continua. El resultado puede ser discreto o continuo, en referencia a las clasificaciones booleanas será un punto en discusión debido a que cada ejemplo se clasifica como verdadero o falso.

**Figura 4**  
*Un árbol de decisión para decidir una opción*



53

Por lo tanto, según (Fu et al., 2023) la **definición conceptual** del algoritmo de Árboles de Decisión es un método de aprendizaje automático supervisado utilizado para la clasificación y regresión. Con este método, las decisiones y sus posibles resultados se representan mediante un modelo de árbol que adopta la forma de una estructura arbórea. En el árbol, cada nodo interno representa un atributo o característica, cada rama una regla de decisión y cada hoja el resultado o previsión.

Dado que el árbol de decisión que se produce se interpreta y muestra de forma tan sencilla, el método de los árboles de decisión es muy conocido por su capacidad de aprendizaje interpretable. También puede utilizarse para resolver problemas de regresión y clasificación, y puede manejar una gran variedad de tipos de variables.

Asimismo, el algoritmo desarrolla dos fases para realizar predicciones:

- **Aprendizaje:** El aprendizaje del algoritmo de Árboles de Decisión implica construir el árbol de decisión utilizando un conjunto de datos de entrenamiento. El proceso se puede dividir en varios pasos:
  - a) **Selección de atributos:** Se eligen los atributos más relevantes o informativos del conjunto de datos para construir el árbol. Esto se puede hacer utilizando diversas técnicas, como la ganancia de información o el índice Gini.
  - b) **División de nodos:** El árbol se construye dividiendo los nodos en función de los atributos seleccionados. Cada división se basa en una regla de decisión, que puede ser una comparación de valores o una condición lógica.
  - c) **Criterio de parada:** La construcción recursiva del árbol se lleva a cabo hasta que se cumple un criterio de parada, como la pureza de clase, el número mínimo de instancias en un nodo o la profundidad máxima.
  - d) **Poda del árbol:** Una vez construido el árbol, puede aplicarse una técnica de poda para evitar el sobreajuste. La poda implica la eliminación de nodos o ramas que no mejoran significativamente la precisión del árbol.
- **Asertividad:** establecer asertividad en la predicción del algoritmo de Árboles de Decisión se refiere a la precisión o exactitud de las predicciones realizadas por el modelo de árbol. Es una medida de qué tan bien el árbol puede clasificar o predecir correctamente nuevas instancias o ejemplos que no se utilizaron durante el proceso de entrenamiento. Por ello asertividad en la predicción puede evaluarse utilizando diversas métricas, dependiendo del tipo de problema, para ello la clasificación incluyen el grado de precisión (Charbuty y Abdulazeez, 2021).

## Naive Bayes

Según Ilić et al. (2022), explica que, aunque los algoritmos típicos de aprendizaje automático se basan en datos fiables, el clasificador Naïve Bayes (NB) funciona bien con datos poco claros, asimismo menciona que el modelo NB es muy popular y efectivo en muchos problemas complejos a pesar de su simplicidad, lo cual es una gran ventaja para su uso; el modelo pertenece a la familia de clasificadores probabilísticos basados en la Teoría de Bayes, cuya característica principal de este clasificador es la suposición de que todas las variables son condicionalmente independientes, razón por la cual se llama 'Naïve'; el uso del algoritmo de aprendizaje automático Naive Bayes está centrado en el desarrollo de problemas de clasificación (Mandal y Jana, 2019).

Villalba (2018) destaca que el modelo es sencillo, robusto y cuya conclusión se deriva directamente de los datos. Su tratamiento se basa en un estadístico bayesiano básico de probabilidad condicional (Barus, 2021), y Villalba destaca que se asume por simplificación que las variables son todas eventos independientes.

Prabhakaran(2018), explica que una técnica probabilística de aprendizaje automático NB puede aplicarse a una serie de aplicaciones de categorización. Son comunes aplicaciones como el análisis de sentimientos, la clasificación de documentos, el filtrado de spam y otras. El nombre proviene del hecho de que se basa en los escritos del reverendo Thomas Bayes. Como se supone que las características que entran en el modelo son independientes entre sí, se utiliza el término "ingenuo". En otras palabras, alterar el valor de una característica no tiene ningún impacto o efecto directo sobre los valores de las demás características que utiliza el algoritmo.

Se trata de un modelo probabilístico con un método fácilmente codificable que permite realizar pronósticos a la velocidad del rayo, incluso en tiempo real. Su rápida capacidad de respuesta lo convierte en una solución altamente escalable, y desde hace tiempo es el algoritmo de referencia para aplicaciones del mundo real en las que es esencial generar respuestas instantáneas a las peticiones de los usuarios.

Mohammed et al. (2017) menciona que los clasificadores bayesianos ingenuos son clasificadores probabilísticos simples basado en aplicar el teorema de Bayes con el supuesto de una (ingenua) independencia entre sus atributos. La siguiente ecuación establece el teorema de Bayes en términos matemáticos:

$$P(A | B) = \frac{P(A)P(B | A)}{P(B)}$$

dónde:

A y B son eventos



$P(A)$  y  $P(B)$  Probabilidad de A y B

$P(A | B)$ , Probabilidad posterior de que se dé A dado B

$P(B | A)$ , Probabilidad posterior de que se dé B dado A

Roman (2019) menciona que este tipo de algoritmo se conoce como una clase específica de algoritmos de clasificación de aprendizaje automático. Se basan en el método de clasificación estadística del "teorema de Bayes". Debido a su suposición simplista (es decir, que todas las variables predictoras son independientes entre sí), estos modelos suelen denominarse algoritmos "ingenuos". Dicho de otro modo, su suposición subyacente es que no existe correlación entre la existencia de una determinada característica en un conjunto de datos y la presencia de cualquier otra característica. A pesar de esta simplificación, estos algoritmos destacan por su facilidad de construcción de modelos y su excelente rendimiento, que puede atribuirse a su sencillez. El experto Brownlee (2018) en inteligencia artificial, destaca la utilización de usar la probabilidad para hacer predicciones. Quizás el ejemplo más utilizado se llama algoritmo Naive Bayes. No solo es fácil de entender, sino que también logra resultados sorprendentemente buenos.

Por ello según Barus (2021), lo **define conceptualmente** al algoritmo Naive Bayes como el método de aprendizaje automático supervisado basado en el teorema de Bayes. Aunque el algoritmo Naive Bayes también puede aplicarse a tareas de regresión, su uso principal se centra en problemas de clasificación. La base de este algoritmo radica en el concepto de independencia condicional de los atributos, el cual significa que, dado un valor de clase específico, cada atributo se considera independiente de los demás. Utilizando la regla de Bayes, Naive Bayes calcula la probabilidad de que una instancia pertenezca a una clase particular. Para lograr esto, el algoritmo estima las probabilidades condicionales y las probabilidades a priori de diversas clases y atributos dentro del conjunto de entrenamiento, permitiéndole asignar con eficacia una clase a cada nueva instancia basada en los datos disponibles

Luego, utiliza estas probabilidades para realizar predicciones sobre nuevas instancias, seleccionando la clase con la probabilidad más alta.

Asimismo, el algoritmo establece indicadores de aprendizaje y asertividad:

- **Aprendizaje**, Utilizando como base un conjunto de datos de entrenamiento, el algoritmo Naive Bayes estima las probabilidades a priori y condicionales de los atributos y las clases. El proceso se puede dividir en los siguientes pasos (Parth Shukla, 2023):
  - a) **Recopilación de datos:** Se recopila un conjunto de datos de entrenamiento que consta de instancias etiquetadas con sus respectivas clases.

- b) Estimación de probabilidades a priori:** Se calculan las probabilidades a priori de las clases, es decir, la probabilidad de que un caso elegido al azar pertenezca a una clase determinada.
  - c) Estimación de probabilidades condicionales:** Se calculan las probabilidades condicionales de los atributos dados los valores de la clase. En el caso del Naive Bayes, se asume independencia condicional entre los atributos, por lo que estas probabilidades se pueden estimar de manera independiente para cada atributo.
  - d) Predicción de nuevas instancias:** Una vez que se han estimado las probabilidades, se utilizan para realizar predicciones sobre nuevas instancias. Se calcula la probabilidad posterior de cada clase para una instancia dada y se selecciona la clase con la probabilidad más alta como la predicción.
- **Asertividad,** Se refiere a la precisión o exactitud de los pronósticos realizados por el modelo. Es una medida de qué tan bien el algoritmo puede clasificar correctamente nuevas instancias o ejemplos que no se utilizaron durante el proceso de entrenamiento. Asertividad en la predicción puede evaluarse utilizando la métrica de precisión dependiendo del problema de clasificación específico y del umbral de decisión seleccionado (Vadlamudi et al., 2023).

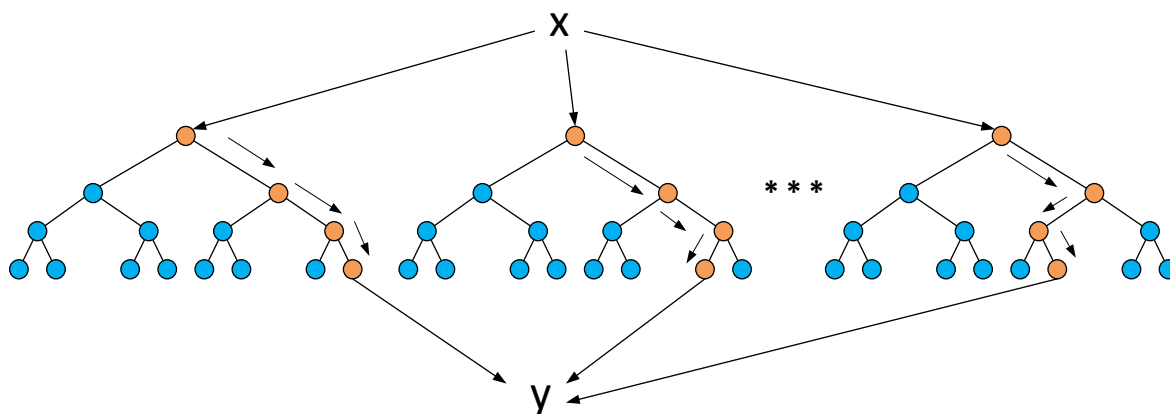
## **Random Forest**

(Pellegrino et al., 2021) explica que Random forest (RF) es un algoritmo ML que se define por un conjunto de métodos de aprendizaje para clasificación, regresión y otras tareas, opera mediante varios árboles de decisión (Pirneskoski et al., 2020), en la fase de entrenamiento y generando las clases (problema de clasificación) o predicción media (problema de regresión) de los árboles individuales; el algoritmo construye cada uno de sus árboles de decisión, entrenándolos a todos con un subconjunto de los datos del problema (Figura 5), elige aleatoriamente N registros de los datos y aprende en múltiples árboles que se construyen y entrenan aleatoriamente de acuerdo con el principio de embolsado (Ong et al., 2022), asimismo Pellegrino asegura que todos los árboles de decisión tengan una experiencia diferente del problema. Una vez entrenados todos los árboles de decisión, RF toma sus decisiones de acuerdo con el problema de clasificación o regresión a resolver votando todos sus árboles de decisión. La aleatoriedad RF se encuentra en el muestreo sobre el que se construyen los árboles

y en la selección de variables sobre las que se realiza la segmentación. Es por eso que RF es una caja negra en muchos casos.

**Figura 5**

*Random Forest*



Nota: Comportamiento de la toma de decisiones

(Schonlau y Zou, 2020) conceptualmente, el algoritmo de aprendizaje automático Random Forest es una técnica de aprendizaje supervisado que puede utilizarse tanto para aplicaciones de regresión como de clasificación. Las predicciones pueden realizarse con mayor precisión y resistencia gracias a su funcionamiento, que combina varios árboles de decisión independientes.

Asimismo, el algoritmo establece dos componentes de desarrollo:

- **Aprendizaje:** se refiere al método mediante el cual el modelo Random Forest aprende de los datos entrenados a fin de establecer pronósticos cada vez más precisos y ampliamente aplicables. Según, Kaur Arora (2022), el algoritmo de aprendizaje automático Random Forest se desarrolla de los siguientes pasos:
  - a) **Preparación de los datos:** Al igual que los demás algoritmos inteligentes, se requiere preparar los datos de entrenamiento, lo cual incluye la selección y transformación de los atributos relevantes y la codificación adecuada de las variables categóricas.
  - b) **Construcción de los árboles de decisión:** Basándose en los datos de entrenamiento, Random Forest crea una colección de árboles de decisión. Para

crear cada árbol se utiliza un método denominado "muestreo bootstrap". Consiste en seleccionar muestras al azar del conjunto de entrenamiento y sustituirlas. Además, en cada nodo del árbol, se selecciona un subconjunto aleatorio de atributos para dividir los datos.

- c) **Votación o promedio de resultados:** Una vez que se han construido todos los árboles, Random Forest realiza la predicción combinando los resultados de cada árbol. En el caso de tareas de clasificación, se utiliza un esquema de votación para determinar la clase final, donde cada árbol emite un voto y la clase más votada se selecciona como predicción. En el caso de tareas de regresión, se toma el promedio de los valores predichos por cada árbol.
- **Asertividad:** Se refiere de la precisión y fiabilidad de las predicciones del modelo en el pronóstico del algoritmo Random Forest. Debido a que Random Forest combina múltiples árboles de decisión, que a su vez aprenden de diferentes subconjuntos de datos y atributos, tiende a tener una mayor capacidad predictiva en comparación con un solo árbol de decisión. La combinación de predicciones de diferentes árboles a través de votación o promedio ayuda a reducir el sesgo y la varianza, lo que puede conducir a un mayor asertividad en las predicciones (Andrade y Freitas, 2023).

## **Definición de términos básicos**

### **Productividad educativa**

Cecilia et al. (2020) menciona que la productividad educativa es la medida en que los estudiantes logran sus metas educativas a corto o largo plazo. Esto se mide comúnmente a través de un examen externo o interno, así como una evaluación continua en forma de pruebas, asignaciones, proyectos, debates y trabajos prácticos y trimestrales.

### **Modelo de Productividad Educativa de Walberg:**

Es un marco teórico que identifica y analiza múltiples variables que afectan el rendimiento académico y la calidad de la educación. Walberg propone que la productividad educativa se ve influida por factores como el tiempo de instrucción, la calidad del personal docente, el clima escolar, la motivación del estudiante, los recursos educativos y otros elementos que interactúan para determinar el éxito educativo (Zambrano, 2022).

### **Capacidad del estudiante**

Destaca la capacidad del estudiante es un contexto amplio de habilidades cognitivas, emocionales y sociales que influyen en la capacidad de aprender y desarrollarse en contextos educativos y más allá. También abarca la disposición para aprender, la motivación, capacidad para resolver problemas, inventiva, flexibilidad y otros rasgos que favorecen el éxito en los entornos de aprendizaje y en la vida en general (Figuerola-Abarzúa et al., 2022).

### **Aptitud del estudiante**

Es un concepto integral que engloba las habilidades, actitudes y disposiciones que contribuyen al desempeño académico y al desarrollo personal del estudiante. Es importante tener en cuenta que la aptitud puede ser moldeada y desarrollada a lo largo del tiempo mediante experiencias educativas, prácticas y esfuerzos individuales (Herrera y Sánchez, 2019).

### **Motivación**

La motivación del estudiante implica la fuerza interna que impulsa al individuo a comprometerse y esforzarse en sus estudios, superar desafíos y persistir en la consecución de sus objetivos educativos. Este concepto reconoce que la motivación no es uniforme para todos los estudiantes y puede variar según factores individuales, contextuales y situacionales (Avila, 2022).

### **Desarrollo**

El desarrollo cognitivo implica cómo los individuos perciben, procesan y utilizan la información del entorno para comprender el mundo que les rodea y adaptarse a él. Este concepto está estrechamente relacionado con la maduración del cerebro, la interacción con el entorno y las experiencias de aprendizaje a lo largo de la vida (Ahmed Khan et al., 2023).

### **Ambiente de aprendizaje**

El término "entorno de aprendizaje" se refiere a un concepto más amplio que incluye factores sociales, emocionales y culturales, además de la disposición física del aula. La relación profesor-alumno, la cooperación entre compañeros, la accesibilidad a los materiales didácticos, la flexibilidad del plan de estudios, el uso de la tecnología y la atmósfera emocional del aula son componentes de este entorno. (Castañón, 2023).

### **Rendimiento académico**

Desde un punto de vista conceptual, el rendimiento académico es más amplio que las notas que se obtienen en las evaluaciones, los exámenes o las valoraciones formales, aunque son indicadores importantes. También incluye elementos adicionales que muestran lo bien que los estudiantes aprovechan las oportunidades de aprendizaje, como la creatividad, la capacidad de

resolución de problemas, la capacidad de análisis y síntesis y la participación activa en clase (Grasso Imig, 2020).

### **Datos**

Es una representación de hechos, conceptos o instrucciones en forma formalizada, apta para su comunicación, interpretación o procesamiento, asimismo es una información creada por un usuario, como documentos, imágenes o grabaciones de sonido (Rigdon, 2016).

### **Dataset o colección de datos**

Es una colección coherente de datos con criterios de selección y estructura interna bien definido (Oxford University, 2016); también es definido como una colección de información relacionada formada por elementos separados que pueden ser tratados como una unidad en el manejo de datos (Rigdon, 2016).

### **Información**

Según Oxford University (2018) la información es un conjunto de datos con significado que es capaz de hacer que la mente humana cambie su opinión sobre el estado actual del mundo real. En ciencia e ingeniería, información es todo lo que contribuye a reducir la incertidumbre del estado de un sistema; en este caso, la incertidumbre suele expresarse de forma objetivamente medible. Es importante mencionar que la información debe distinguirse de cualquier medio que sea capaz de transportarla. Un medio físico (como un disco magnético) puede transportar un medio lógico (datos, como símbolos binarios o de texto). El contenido de información de cualquier objeto físico, o datos lógicos, no puede medirse ni discutirse hasta que se sepa qué rango de posibilidades existía antes y después de que fueran recibidos.

La información radica en la reducción de la incertidumbre resultante de la recepción de los objetos o datos, y no en el tamaño o complejidad de los objetos o datos en sí. Las preguntas sobre la forma, la función y la importancia semántica de los datos solo son relevantes para la información en la medida en que contribuyan a la reducción de la incertidumbre. Si un memorándum idéntico se recibe dos veces, no transmite el doble de información que la primera vez: la segunda no transmite ninguna información, a menos que, por acuerdo previo, el número de ocurrencias se considere significativo.

### **Patrón**

Laplante (2001) menciona que es una regla que especifica clases de equivalencia de secuencias textuales. Un patrón a menudo se expresa en una notación como una expresión regular o alguna notación equivalente. Los lenguajes como SNOBOL tenían especificaciones de patrones elaborados que incluían reglas para volver a intentar una coincidencia después de un

intento fallido ("reglas de retroceso"). Los lenguajes modernos que incluyen patrones de coincidencia de cadenas como operaciones primitivas incluyen awk y Perl.

### **Algoritmo**

Según, Oxford University (2018) un algoritmo es una colección predeterminada de pautas o directivas creadas para completar una tarea, como completar un cálculo, en una cantidad limitada de pasos. Un componente clave de la programación automatizada es la notación formal de algoritmos; muchas ideas que se aplican en programas se aplican en algoritmos y en forma contraria. Cuando un algoritmo se puede calcular con éxito, se dice que es eficiente. La base de la teoría de algoritmos es la investigación de la existencia de algoritmos capaces de calcular determinadas cantidades. Puede resultar difícil demostrar la solidez de un algoritmo o incluso determinar su objetivo preciso. La validación de algoritmos, un procedimiento que confirma que un algoritmo cumple sus funciones, se utiliza con frecuencia en la práctica. Esto involucra poner a prueba el algoritmo en una serie de escenarios para garantizar su funcionamiento adecuado en cada uno de ellos. Se puede estar seguro de que el algoritmo es fiable eligiendo un conjunto de pruebas adecuado.

### **Machine learning (ml)**

Raynor (999) define a Machine Learning como la capacidad de un programa para adquirir o desarrollar nuevos conocimientos o habilidades. El estudio de Machine Learning se centra en el desarrollo de métodos de cálculo para descubrir nuevos conocimientos a partir de datos.

BBVA (2019) define como el área de la inteligencia artificial que se ocupa del aprendizaje automático sin necesidad de programación especial. una característica crucial que permite a los sistemas reconocer patrones en los datos para poder generar predicciones. Numerosas aplicaciones, como los contenidos sugeridos de Netflix o Spotify, las respuestas inteligentes de Gmail o el reconocimiento de voz de Siri y Alexa, utilizan esta técnica.

### **Aprendizaje supervisado**

Un subcampo de la inteligencia artificial y el aprendizaje automático se denomina aprendizaje supervisado o aprendizaje automático supervisado. Se distingue por el uso de conjuntos de datos etiquetados para entrenar algoritmos de clasificación precisa de datos o predicción de resultados. A lo largo de este procedimiento, el modelo aprende de los datos que recibe y modifica sus ponderaciones según sea necesario para garantizar un buen ajuste del modelo. Las organizaciones pueden utilizar el aprendizaje supervisado para resolver una serie de problemas complejos del mundo real, como clasificar automáticamente los correos electrónicos no solicitados en una carpeta diferente de la bandeja de entrada (IBM Cloud Education, 2020).

## **Aprendizaje no supervisado**

El aprendizaje no supervisado, usualmente acreditado como aprendizaje automático no supervisado, utiliza algoritmos de aprendizaje automático para agrupar y examinar conjuntos de datos no etiquetados. Sin la ayuda de una persona, estos algoritmos pueden encontrar patrones ocultos o agrupar datos. Por su capacidad para identificar patrones en los datos, es perfecto para actividades como la ejecución de estrategias de venta cruzada, el reconocimiento de imágenes, la segmentación de clientes y el análisis exploratorio de datos (IBM Cloud Education, 2020).

## **Clasificación**

Predecir a qué clase pertenece un conjunto concreto de puntos de datos es el proceso de clasificación. Estas clases también pueden denominarse etiquetas, categorías u objetivos. El proceso de aproximación de una función de mapeo ( $f$ ) que conecta variables de salida discretas ( $Y$ ) con variables de entrada ( $X$ ) se conoce como modelización de clasificación predictiva. (Asiri, 2018).

Laplante (2001) menciona que es el proceso de encontrar reglas de clasificación, adicionalmente menciona que es un área de minería de datos que intenta predecir la categoría de datos categóricos mediante la construcción de un modelo basado en algunas variables predictoras.

## **Referencias**

- Abid, N. y Ali, R. y Akhter, M. (2021). Exploring gender-based difference towards academic enablers scales among secondary school students of Pakistan. *Psychology in the Schools*, 58(7), 1380–1398. <https://doi.org/10.1002/pits.22538>
- Ahmed Khan, Z. y Adnan, J. y Adnan Raza, S. (2023). *Cognitive Learning Theory and Development: Higher Education Case Study*. <https://doi.org/10.5772/intechopen.110629>
- Al-Ani, O. y Das, S. (2022). Reinforcement Learning: Theory and Applications in HEMS. *MATHEMATICS & COMPUTER SCIENCE, Artificial Intelligence & Robotics*.
- Andrade, M. S. y Freitas, J. C. de. (2023). Analysis of Performance Metrics on the Conjunction of Intrusions in IEEE 802.11 Networks with Machine Learning at Hospital N.S.C. In *Connecting Expertise Multidisciplinary Development For The Future*. Seven Editora. <https://doi.org/10.56238/Connexpemultidisdevolpfut-116>
- Avila, Á. (2022). *La motivación en los estudiantes: 3 Claves para aumentarla*. <https://www.enriccorberainstitute.com/blog/motivacion-en-los-estudiantes/>



- Bagnato Juan Ignacio. (2020). *Aprende Machine Learning en Español Teoría + Práctica Python*. 1–184.
- Barus, S. P. (2021). Implementation of Naïve Bayes Classifier-based Machine Learning to Predict and Classify New Students at Matana University. *Journal of Physics: Conference Series*, 1842(1), 012008. <https://doi.org/10.1088/1742-6596/1842/1/012008>
- BBVA. (2019). “Machine learning”: ¿qué es y cómo funciona?
- Brownlee, J. (2018). *Machine Learning Algorithms From Scratch* (v.1.8, Ed.).
- Carroll, J. B. (1989). The Carroll Model: A 25-Year Retrospective and Prospective View. *Educational Researcher*, 18(1), 26. <https://doi.org/10.2307/1176007>
- Caselli Gismondi, H. E. (2021). *Modelo predictivo basado en machine learning como soporte para el seguimiento académico del estudiante universitario* [Universidad Nacional del Santa].  
<http://repositorio.uns.edu.pe/bitstream/handle/UNS/3804/52337.pdf?sequence=5&isAllowed=y>
- Castañón, E. R. (2023). *¿Qué es un ambiente de aprendizaje según autores?*  
<https://www.centrobanamex.com.mx/que-es-un-ambiente-de-aprendizaje-segun-autores/>
- Cecilia, O. N. y Bernedette, C.-U. y Emmanuel, A. E. y Hope, A. N. (2020). Enhancing students academic performance in Chemistry by using kitchen resources in Ikom, Calabar. *Educational Research and Reviews*, 15(1), 19–26.  
<https://doi.org/10.5897/ERR2019.3810>
- Charbuty, B. y Abdulazeez, A. (2021). Classification Based on Decision Tree Algorithm for Machine Learning. *Journal of Applied Science and Technology Trends*, 2(01), 20–28.  
<https://doi.org/10.38094/jastt20165>
- Díaz-Landa, B. y Meleán-Romero, R. y Marín-Rodríguez, W. (2021). Rendimiento académico de estudiantes en Educación Superior: predicciones de factores influyentes a partir de árboles de decisión. *Telos Revista de Estudios Interdisciplinarios En Ciencias Sociales*, 23(3), 616–639. <https://doi.org/10.36390/telos233.08>
- Dong, Y. y Hou, J. y Zhang, N. y Zhang, M. (2020). Research on How Human Intelligence, Consciousness, and Cognitive Computing Affect the Development of Artificial Intelligence. *Complexity*, 2020, 1–10. <https://doi.org/10.1155/2020/1680845>
- Esteso, A. y Peidro, D. y Mula, J. y Díaz-Madroñero, M. (2022). Reinforcement learning applied to production planning and control. *International Journal of Production Research*, 1–18.  
<https://doi.org/10.1080/00207543.2022.2104180>

- Figuerola-Abarzúa, C. y Meza-Vásquez, S. y Estrada-Lagos, R. (2022). Desarrollo de capacidad argumentativa en estudiantes universitarios, mediante uso del debate como estrategia didáctica. *South Florida Journal of Development*, 3(6), 6328–6346. <https://doi.org/10.46932/sfjdv3n6-001>
- Fu, M. y Zhang, C. y Hu, C. y Wu, T. y Dong, J. y Zhu, L. (2023). Achieving Verifiable Decision Tree Prediction on Hybrid Blockchains. *Entropy*, 25(7), 1058. <https://doi.org/10.3390/e25071058>
- Géron, A. (2019). *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow Concepts, Tools, and Techniques to Build Intelligent System* (Ni. Tache, Ed.; 2nd ed.).
- Gogus, A. (2012). Bloom's Taxonomy of Learning Objectives. In *Encyclopedia of the Sciences of Learning* (pp. 469–473). Springer US. [https://doi.org/10.1007/978-1-4419-1428-6\\_141](https://doi.org/10.1007/978-1-4419-1428-6_141)
- Grasso Imig, P. (2020). *Rendimiento académico: un recorrido conceptual que aproxima a una definición unificada para el ámbito superior*. [https://fh.mdp.edu.ar/revistas/index.php/r\\_educ/article/view/4165](https://fh.mdp.edu.ar/revistas/index.php/r_educ/article/view/4165)
- Guill, K. y Ömeroğulları, M. y Köller, O. (2021). Intensity and content of private tutoring lessons during German secondary schooling: effects on students' grades and test achievement. *European Journal of Psychology of Education*. <https://doi.org/10.1007/s10212-021-00581-x>
- Hajjej, F. y Alohali, M. A. y Badr, M. y Rahman, M. A. (2022). A Comparison of Decision Tree Algorithms in the Assessment of Biomedical Data. *BioMed Research International*, 2022, 1–9. <https://doi.org/10.1155/2022/9449497>
- Herrera Acosta, C. E. y Sánchez Pinedo, L. D. (2019). Metodología De Aprendizaje Y Sistema De Evaluación Para Alcanzar Resultados En El Proceso Educativo. *Evaluación de La Calidad Educativa*. <https://www.igobernanza.org/index.php/IGOB/Contacto>
- IBM Cloud Education. (2020). *¿Qué es el aprendizaje supervisado?* IBM. <https://www.ibm.com/es-es/think/topics/supervised-learning>
- Ilić, M. y Srdjević, Z. y Srdjević, B. (2022). Water quality prediction based on Naïve Bayes algorithm. *Water Science and Technology*, 85(4), 1027–1039. <https://doi.org/10.2166/wst.2022.006>
- Janiesch, C. y Zschech, P. y Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3), 685–695. <https://doi.org/10.1007/s12525-021-00475-2>
- Jiang, T. y Gradus, J. L. y Rosellini, A. J. (2020). Supervised Machine Learning: A Brief Primer. *Behavior Therapy*, 51(5), 675–687. <https://doi.org/10.1016/j.beth.2020.05.002>

- Jovel, J. y Greiner, R. (2021). An Introduction to Machine Learning Approaches for Biomedical Research. *Frontiers in Medicine*, 8. <https://doi.org/10.3389/fmed.2021.771607>
- Kaun, C. y Jhanjhi, N. Z. y Goh, W. W. y Sukumaran, S. (2021). Implementation of Decision Tree Algorithm to Classify Knowledge Quality in a Knowledge Intensive System. *MATEC Web of Conferences*, 335, 04002. <https://doi.org/10.1051/mateconf/202133504002>
- Kaur Arora, S. (2022). *Top Steps To Learn Naive Bayes Algorithm*. <https://hackr.io/blog/top-steps-to-learn-naive-bayes-algorithm>
- KOÇOĞLU, F. Ö. (2022). Research on the success of unsupervised learning algorithms in indoor location prediction. *International Advanced Researches and Engineering Journal*, 6(2 (under construction)), 148–153. <https://doi.org/10.35860/iarej.1096573>
- Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. In *Theory Into Practice* (Vol. 41, Issue 4). [https://doi.org/10.1207/s15430421tip4104\\_2](https://doi.org/10.1207/s15430421tip4104_2)
- Lampropoulos, A. S. y Tsihrintzis, G. A. (2015). *Machine Learning Paradigms* (Vol. 92). Springer International Publishing. <https://doi.org/10.1007/978-3-319-19135-5>
- Laplante, P. A. (2001). *Dictionary of computer science, engineering, and technology*. [www.crcpress.com](http://www.crcpress.com)
- Lee, J. y Warner, E. y Shaikhouni, S. y Bitzer, M. y Kretzler, M. y Gipson, D. y Pennathur, S. y Bellovich, K. y Bhat, Z. y Gadegbeku, C. y Massengill, S. y Perumal, K. y Saha, J. y Yang, Y. y Luo, J. y Zhang, X. y Mariani, L. y Hodgins, J. B. y Rao, A. (2022). Unsupervised machine learning for identifying important visual features through bag-of-words using histopathology data from chronic kidney disease. *Scientific Reports*, 12(1), 4832. <https://doi.org/10.1038/s41598-022-08974-8>
- Mandal, L. y Jana, N. D. (2019). A Comparative Study of Naive Bayes and k-NN Algorithm for Multi-class Drug Molecule Classification. *2019 IEEE 16th India Council International Conference (INDICON)*, 1–4. <https://doi.org/10.1109/INDICON47234.2019.9029095>
- Mazana, M. Y. y Montero, C. S. y Casmir, R. O. (2019). Investigating Students' Attitude towards Learning Mathematics. *International Electronic Journal of Mathematics Education*, 14(1). <https://doi.org/10.29333/iejme/3997>
- Mienye, I. D. y Sun, Y. y Wang, Z. (2019). Prediction performance of improved decision tree-based algorithms: a review. *Procedia Manufacturing*, 35, 698–703. <https://doi.org/10.1016/j.promfg.2019.06.011>
- Mitchell, T. M. (1997). *Machine learning*. McGraw-Hill. <https://www.cs.cmu.edu/~tom/files/MachineLearningTomMitchell.pdf>

- Mohammed, M. y Khan, M. B. y Mohammed Bashier, E. B. (2017). *Machine Learning Algorithms and Applications*.
- Morales Hernández, M. Á. y González Camacho, J. M. y Robles Vásquez, H. y Del Valle Paniagua, D. H. y Durán Moreno, J. R. (2022). Algoritmos de aprendizaje automático para la predicción del logro académico. *RIDE Revista Iberoamericana Para La Investigación y El Desarrollo Educativo*, 12(24). <https://doi.org/10.23913/ride.v12i24.1180>
- Münch, M. y Raab, C. y Biehl, M. y Schleif, F.-M. (2020). Data-Driven Supervised Learning for Life Science Data. *Frontiers in Applied Mathematics and Statistics*, 6. <https://doi.org/10.3389/fams.2020.553000>
- Nedeva, V. y Pehlivanova, T. (2021). Students' Performance Analyses Using Machine Learning Algorithms in WEKA. *IOP Conference Series: Materials Science and Engineering*, 1031(1), 012061. <https://doi.org/10.1088/1757-899X/1031/1/012061>
- Ong, A. K. S. y Chuenyindee, T. y Prasetyo, Y. T. y Nadlifatin, R. y Persada, S. F. y Gumasing, Ma. J. J. y German, J. D. y Robas, K. P. E. y Young, M. N. y Sittiwatethanasiri, T. (2022). Utilization of Random Forest and Deep Learning Neural Network for Predicting Factors Affecting Perceived Usability of a COVID-19 Contact Tracing Mobile Application in Thailand “ThaiChana.” *International Journal of Environmental Research and Public Health*, 19(10), 6111. <https://doi.org/10.3390/ijerph19106111>
- Oxford University. (2016). *A Dictionary of Computer Science*.
- Parhizkar, A. y Tejeddin, G. y Khatibi, T. (2023). Student performance prediction using datamining classification algorithms: Evaluating generalizability of models from geographical aspect. *Education and Information Technologies*, 28(11), 14167–14185. <https://doi.org/10.1007/s10639-022-11560-0>
- Parth Shukla. (2023). *Naive Bayes Algorithms: A Complete Guide for Beginners*. <https://www.analyticsvidhya.com/blog/2023/01/naive-bayes-algorithms-a-complete-guide-for-beginners/>
- Peconcillo Jr, L. B. y D. Peteros, E. y O. Mamites, I. y T. Sanchez, D. y L. Tenerife, J. J. y L. Suson, R. (2020). Structuring Determinants to Level Up Students Performance. *International Journal of Education and Practice*, 8(4), 638–651. <https://doi.org/10.18488/journal.61.2020.84.638.651>
- Pellegrino, E. y Jacques, C. y Beaufils, N. y Nanni, I. y Carliz, A. y Metellus, P. y Ouafik, L. (2021). Machine learning random forest for predicting oncosomatic variant NGS analysis. *Scientific Reports*, 11(1), 21820. <https://doi.org/10.1038/s41598-021-01253-y>

- Pirneskoski, J. y Tamminen, J. y Kallonen, A. y Nurmi, J. y Kuisma, M. y Olkkola, K. T. y Hoppu, S. (2020). Random forest machine learning method outperforms prehospital National Early Warning Score for predicting one-day mortality: A retrospective study. *Resuscitation Plus*, 4, 100046. <https://doi.org/10.1016/j.resplu.2020.100046>
- Prabhakaran, S. (2018). *How Naive Bayes Algorithm Works?* Machine Learning. <https://www.machinelearningplus.com/predictive-modeling/how-naive-bayes-algorithm-works-with-example-and-full-code/>
- Pugliese, R. y Regondi, S. y Marini, R. (2021). Machine learning-based approach: global trends, research directions, and regulatory standpoints. *Data Science and Management*, 4, 19–29. <https://doi.org/10.1016/j.dsm.2021.12.002>
- Quiñones, L. y Quiñones, Y. L. (2020). Rendimiento académico empleando minería de datos. *Espacios*, 41(44), 277–285. <https://doi.org/10.48082/espacios-a20v41n44p17>
- Raynor, W. J. (1999). *The International Dictionary of Artificial Intelligence*.
- Rigdon, J. C. (2016). *Dictionary of Computer and Internet Terms* (Microsoft Corporation, Ed.; 1st Edition). <http://www.wordsrus.info>
- Roman, V. (2019). *Algoritmos Naive Bayes: Fundamentos e Implementación*. Ciencia de Datos. <https://medium.com/datos-y-ciencia/algoritmos-naive-bayes-fudamentos-e-implementaci%C3%B3n-4bcb24b307f>
- Rosales Sánchez, E. M. y Rodríguez Ortega, P. G. y Romero Ariza, M. (2020). Conocimiento, demanda cognitiva y contextos en la evaluación de la alfabetización científica en PISA. *Revista Eureka Sobre Enseñanza y Divulgación de Las Ciencias*, 17(2), 1–22. [https://doi.org/10.25267/Rev\\_Eureka\\_ensen\\_divulg\\_cienc.2020.v17.i2.2302](https://doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2020.v17.i2.2302)
- Russell, S. y Norvig, P. (2008). *Inteligencia Artificial Un Enfoque Moderno* (D. F. Aragón, Ed.; 2nd ed.).
- Sauce, B. y Liebherr, M. y Judd, N. y Klingberg, T. (2022). The impact of digital media on children's intelligence while controlling for genetic differences in cognition and socioeconomic background. *Scientific Reports*, 12(1), 7720. <https://doi.org/10.1038/s41598-022-11341-2>
- Schonlau, M. y Zou, R. Y. (2020). The random forest algorithm for statistical learning. *The Stata Journal: Promoting Communications on Statistics and Stata*, 20(1), 3–29. <https://doi.org/10.1177/1536867X20909688>
- Seel, N. M. (2012a). Bloom's Model of School Learning. In *Encyclopedia of the Sciences of Learning* (pp. 466–469). Springer US. [https://doi.org/10.1007/978-1-4419-1428-6\\_979](https://doi.org/10.1007/978-1-4419-1428-6_979)

- Seel, N. M. (2012b). Carroll's Model of School Learning. In *Encyclopedia of the Sciences of Learning* (pp. 501–503). Springer US. [https://doi.org/10.1007/978-1-4419-1428-6\\_980](https://doi.org/10.1007/978-1-4419-1428-6_980)
- Sepúlveda, A. y Minte, A. y Díaz-Levicoy, D. (2020). Caracterización de preguntas en libros de texto de Ciencias Naturales en Educación Primaria chilena. *Educação e Pesquisa*, 46. <https://doi.org/10.1590/s1678-4634202046224118>
- Serumena D. y Utan F. y Poernomo M. (2021). The Effectiveness of Social Media as An Online Learning Pattern in Improving the 3 Domains of Student Intellectual Ability During the Pandemic (Covid-19). *Advances in Engineering Research*.
- Simplilearn. (2022). *The Complete Guide to Machine Learning Steps*. <https://www.simplilearn.com/tutorials/machine-learning-tutorial/machine-learning-steps>
- Singh, V. y Chen, S.-S. y Singhanian, M. y Nanavati, B. y kar, A. kumar y Gupta, A. (2022). How are reinforcement learning and deep learning algorithms used for big data based decision making in financial industries—A review and research agenda. *International Journal of Information Management Data Insights*, 2(2), 100094. <https://doi.org/10.1016/j.jjime.2022.100094>
- Stern, M. y Arinze, C. y Perez, L. y Palmer, S. E. y Murugan, A. (2020). Supervised learning through physical changes in a mechanical system. *Proceedings of the National Academy of Sciences*, 117(26), 14843–14850. <https://doi.org/10.1073/pnas.2000807117>
- Tai, M.-T. (2020). The impact of artificial intelligence on human society and bioethics. *Tzu Chi Medical Journal*, 32(4), 339. [https://doi.org/10.4103/tcmj.tcmj\\_71\\_20](https://doi.org/10.4103/tcmj.tcmj_71_20)
- Tebani, A. y Bekri, S. (2020). High-throughput omics in the precision medicine ecosystem. In *Precision Medicine for Investigators, Practitioners and Providers* (pp. 19–31). Elsevier. <https://doi.org/10.1016/B978-0-12-819178-1.00003-4>
- Vadlamudi, P. S. y Gunasekaran, M. y Nagalakshmi, T. J. (2023). An Analysis of the Effectiveness of the Naive Bayes Algorithm and the Support Vector Machine for Detecting Fake News on Social Media. *2023 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE)*, 726–731. <https://doi.org/10.1109/IITCEE57236.2023.10090978>
- Vigo, R. y Zeigler, D. E. y Wimsatt, J. (2022). Uncharted Aspects of Human Intelligence in Knowledge-Based “Intelligent” Systems. *Philosophies*, 7(3), 46. <https://doi.org/10.3390/philosophies7030046>
- Villalba, F. (2018). Naive Bayes- clasificación bayesiano ingenuo. In *Aprendizaje supervisado en R*. <https://fervilber.github.io/Aprendizaje-supervisado-en-R/ingenuo.html>

- Vostroknutov, A. y Polonio, L. y Coricelli, G. (2018). The Role of Intelligence in Social Learning. *Scientific Reports*, 8(1), 6896. <https://doi.org/10.1038/s41598-018-25289-9>
- Walberg, H. J. (1984). *Improving the Productivity of America's Schools*.
- Wang, P. (2019). On Defining Artificial Intelligence. *Journal of Artificial General Intelligence*, 10(2), 1–37. <https://doi.org/10.2478/jagi-2019-0002>
- Xiang, X. y Foo, S. y Zang, H. (2021). Recent Advances in Deep Reinforcement Learning Applications for Solving Partially Observable Markov Decision Processes (POMDP) Problems Part 2—Applications in Transportation, Industries, Communications and Networking and More Topics. *Machine Learning and Knowledge Extraction*, 3(4), 863–878. <https://doi.org/10.3390/make3040043>
- Yağcı, M. (2022). Educational data mining: prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments*, 9(1), 11. <https://doi.org/10.1186/s40561-022-00192-z>
- Yildiz, M. y Börekci, C. (2020). Predicting Academic Achievement with Machine Learning Algorithms. *Journal of Educational Technology and Online Learning*. <https://doi.org/10.31681/jetol.773206>
- Zambrano Aranda, G. (2022). Análisis de la satisfacción con respecto a una maestría de una universidad de Lima desde la perspectiva de sus egresados. *Revista Educación y Sociedad*, 3(5), 9–22. <https://doi.org/10.53940/reys.v3i5.89>
- Zhao, W. (2022). Inspired, but not mimicking: a conversation between artificial intelligence and human intelligence. *National Science Review*, 9(6). <https://doi.org/10.1093/nsr/nwac068>
- Zheng, Y. (2019). Identification of microRNAs From Small RNA Sequencing Profiles. In *Computational Non-coding RNA Biology* (pp. 35–82). Elsevier. <https://doi.org/10.1016/B978-0-12-814365-0.00012-9>

# CAPÍTULO 4

## Hipótesis, variables y metodología de la investigación

### Hypothesis, variables and research methodology

DOI: <https://doi.org/10.71112/w474fh82>

#### Resumen

Este capítulo detalla la metodología de investigación para analizar la relación entre la productividad educativa y los algoritmos de aprendizaje automático en estudiantes universitarios. Se plantean hipótesis específicas que vinculan dimensiones como aptitud, instrucción, medio ambiente y aprendizaje con dichos algoritmos. El estudio es de tipo cuantitativo, no experimental, transversal y correlacional, utilizando un cuestionario y una lista de cotejo para la recolección de datos. La muestra consta de 90 estudiantes, y los datos se analizan con SPSS v26, incluyendo la implementación de algoritmos de machine learning.

**Palabras clave:** Metodología Cuantitativa, Hipótesis, Productividad Educativa, Aprendizaje Automático, Diseño Correlacional, SPSS.

#### Abstract

This chapter outlines the research methodology to analyze the relationship between educational productivity and machine learning algorithms in university students. Specific hypotheses link dimensions such as aptitude, instruction, environment, and learning with these algorithms. The study is quantitative, non-experimental, cross-sectional, and correlational, using a questionnaire and a checklist for data collection. The sample consists of 90 students, and data is analyzed with SPSS v26, including the implementation of machine learning algorithms.

**Keywords:** Quantitative Methodology, Hypotheses, Educational Productivity, Machine Learning, Correlational Design, SPSS.

#### Hipótesis general

- Existe relación significativa de la productividad educativa con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023

#### Hipótesis específicas



- Existe relación significativa en la dimensión aptitud con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023
- Existe relación significativa en la dimensión instrucción con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023
- Existe relación significativa en la dimensión medio ambiente con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023
- Existe relación significativa en la dimensión de aprendizaje con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023

## Definición conceptual y operacional de variables

### Productividad educativa

Es el proceso de aprendizaje ligado elementos característicos de aptitud, instrucción y ambiente del entorno social que influyen en la complejidad del aprendizaje humano a fin de conocer o predecir el rendimiento final del estudiante. (Walberg 1981, como se citó en Jin et al., 2022; Maina et al., 2021). Las dimensiones son aptitud, instrucción, medio ambiente y aprendizaje.

### Algoritmos de aprendizaje automatico. –

Son métodos basados en modelos matemáticos que utilizan las computadoras a fin facilitar su aprendizaje o descubrir patrones de datos sin ser programadas explícitamente; es decir es el campo de la informática en el las máquinas se pueden programar a sí mismas. (Sarker, 2021; Palanichamy, 2019). Las dimensiones son algoritmos de árboles de decisión, Naive Bayes, Random Forest y aprendizaje automático.

### Cuadro de operacionalización de variables

Variable	Definición conceptual	Definición operacional	Dimensi ón	Indicadore s	Íte ms	Escal a
Variable 1 Productivi dad educativa	Es el proceso de aprendizaje ligado elementos característicos de aptitud,	Consiste desarrollar una evaluación holística de la teoría a partir de los indicadores involucrados en cada	Aptitud	• Capacida d • Motivació n • Desarroll o	1, 2, 3, 4, 5, 6	Ordi nal

	instrucción y ambiente del entorno social que influyen en la complejidad del aprendizaje humano a fin de predecir el rendimiento final del estudiante.(Walberg 1981, como se citó en Jin et al., 2022; Maina et al., 2021).	dimensión de aptitud, medio ambiente, instrucción a fin de analizar y conocer el logro del aprendizaje. (Jin et al., 2022)	Medio ambiente	• Hogar • Entorno comunitario • Medios de comunicación	7, 8, 9, 10, 11, 12, 13	Ordinal
			Instrucción	• Cantidad • Calidad	14, 15, 16, 17, 18	Ordinal
			Aprendizaje	• Logro	19	Ordinal
Variable 2	Son métodos basados en modelos matemáticos que utilizan las computadoras a fin facilitar su aprendizaje o descubrir patrones de datos sin ser programadas explícitamente	Los algoritmos de aprendizaje automático tienen la capacidad de realizar el reconocimiento de patrones, la clasificación y la predicción de los datos aprendiendo de los datos existentes (conjunto de entrenamiento).(Palanichamy, 2019)	Algoritmo de Arboles de decisión	•Aprendizaje •Asertividad	01, 02, 03, 04, 05, 06	Ordinal
Algoritmo de aprendizaje automático			Algoritmo de Naive Bayes	•Aprendizaje •Asertividad	07, 08, 09, 010, 011, 012	Ordinal

; es decir es el campo de la informática en el las máquinas se pueden programar a sí mismas. (Sarker, 2021; Palanichamy, 2019).	Algoritmo de Random Forest	•Aprendizaje •Asertividad	013 , 014 , 015, 016 , 017, 018	Ordinal
	Aprendizaje Automático	•Predicción	019	Ordinal

## Metodología de la investigación

### Enfoque, tipo y nivel de investigación

#### Enfoque de investigación

Debido a la información obtenida, la investigación utilizó una técnica cuantitativa y conforme a lo establecido por Hernández Sampieri et al. (2014) representa a un conjunto de procesos rigurosos secuenciales, planificados y probados; asimismo las variables son medibles en un determinado contexto, los cuales son analizados estadísticamente, en donde se extraen conclusiones respecto a la hipótesis.

#### Tipo de investigación

El tipo de investigación que se aplicó al presente trabajo es de tipo **Básica**, cuyo propósito según (Hernandez, Fernandez, & Baptista, 2014) busca producir conocimiento y teorías. La investigación básica se apoya dentro de un contexto teórico, esta forma de investigación emplea cuidadosamente el procesamiento del muestreo a fin de extender sus hallazgos más allá de las situaciones estudiadas (Tamayo, 2004).

#### Nivel de investigación

De acuerdo al escenario de la investigación, reúne por su nivel las características de un estudio descriptivo – correlacional; descriptivo porque según (Sampieri & Mendoza, 2018) el nivel descriptivo tiene como finalidad especificar propiedades, fenómenos o hechos en un determinado contexto, en donde se definen y miden las variables y características del planteamiento, se cuantifican y muestran con precisión las dimensiones del fenómeno,

contexto o situación; por otro lado (Ñaupas, Valdivia, Palacios, & Romero, 2018) fortalece lo mencionado, al nivel descriptivo como la recopilación de datos a partir de las características, aspectos o dimensiones o dimensiones, factores, clasificación de objetos o procesos naturales o sociales; asimismo (Gay, 1996, como se citó en Ñaupas et.al., 2013) establece que la investigación descriptiva comprende en una colección de datos para probar la hipótesis o responder preguntas de la investigación; asimismo el estudio descriptivo determina e informa los modos de ser de los objetos.

Por otro lado, es correlacional porque según (Hernandez, Fernandez, & Baptista, 2014) tiene como finalidad establecer la relación existente entre dos o más conceptos, categorías o variables en una muestra particular, por lo general se ubican en un vínculo de estudio entre tres, cuatro o más variables, así mismo (Sampieri & Mendoza, 2018) establece que permiten cierto grado de predicción.

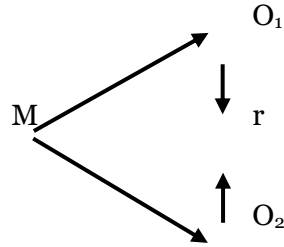
## **Métodos y diseño de investigación.**

### **Métodos de investigación**

En este trabajo se aplicó el enfoque de investigación hipotético-deductivo, que parte de la elaboración de hipótesis. Antes de crear una teoría, este método trata de probar y refutar las hipótesis ofrecidas para asegurarse de que las conclusiones básicas se ven confirmadas y cuestionadas por los hechos (Bernal, 2016). Además, la presente investigación reconoce un corpus teórico a partir del cual se elaboran hipótesis que posteriormente se verifican empíricamente. Esto sugiere que las teorías se derivan de un marco teórico más amplio en lugar de crearse de forma independiente (Maletta, 2000). Lo que plantea la presente investigación en el método de investigación es establecer la verdad o falsedad de la hipótesis de las consecuencias observadas y que llega a unas conclusiones a través de un procedimiento de inferencia o cálculo formal.

### **Diseño de la investigación**

La investigación fue un diseño no experimental, debido a que no existe manipulación de variables, los sujetos de estudios son analizados en un contexto natural, sin alteración (Arias & Covinos, 2021), también está definida como la investigación sin manipular deliberadamente las variables, esto significa que no variamos intencionalmente las variables independientes, para ver los efectos sobre las otras variables, lo que se hace es observar el fenómeno en su contexto natural. (Hernandez, Fernandez, & Baptista, 2014); asimismo es transversal, porque recoge datos en un solo momento.



*Donde:*

- M = Muestra  
 O<sub>1</sub> = Observación de la V1  
 O<sub>2</sub> = Observación de la V2  
 r = Correlación entre dichas variables

## **Población y muestra de la investigación**

### **Población**

La población del presente trabajo de investigación está conformada por estudiantes EPISI – UNAM, del año académico 2023, fue de 116 alumnos, conforme a la información.

### **Muestra**

Dado que la selección de los elementos se basa en la probabilidad de que cualquier elemento sea seleccionado, la muestra se creó utilizando un muestreo aleatorio simple.

El tamaño de la muestra según fórmula:

$$n = \frac{Z^2 p \cdot q N}{e^2 (N - 1) + Z^2 p \cdot q}$$

Dónde:

N = Tamaño de la población = 116

Z = Nivel de confianza = 1.96

e = Margen de error permitido 5%

p = Probabilidad de ocurrencia del evento = 0.5

q = Probabilidad de no ocurrencia del evento = 0.5

n = Tamaño óptimo de la muestra.

Aplicando valores a la formula:

$$n = \frac{(1.96)^2 (0.5) (116)}{0.05^2 (116 - 1) + (1.96)^2 (0.5) (0.5)}$$

Para efectos de la presente investigación se ha realizado un muestreo aleatorio simple, en donde aplicando la formula estableció el tamaño a 90 alumnos de EPISI de UNAM.

## **Técnicas e instrumentos de recolección de datos**

### **Técnicas**

El método empleado en este estudio para investigar la variable 1: Productividad educativa es la encuesta y para la variable 2: Algoritmos de aprendizaje automático será la observación, para ello Carrasco (2006) señala que el enfoque es el proceso que facilita la recogida y compilación de datos a partir de documentos relacionados con el tema y el objetivo del estudio. Además, se hace hincapié en lo crucial que es elegir cuidadosamente la documentación para asegurarse de que respalda los objetivos del proyecto de investigación.

### **Instrumentos**

Según Carrasco (2006), es posible llevar a cabo una investigación científica -un proceso metódico de indagación y búsqueda de información sobre los hechos y fenómenos de la realidad- mediante el uso de instrumentos, que también son cruciales para la recopilación de datos.

El presente trabajo de investigación utilizo para la variable 1: Productividad educativa, el cuestionario y para la variable 2: Algoritmos de aprendizaje automático será lista de cotejo para la, el mismo que permitió recolectar datos significativos para su análisis.

### **Validez**

El Informe de validez del instrumento requirió la opinión de expertos. Para ello, se envió el cuestionario a cinco (5) jueces profesionales, todos ellos doctores.

Es importante mencionar que la validez de los instrumentos requiere opiniones de juicio de expertos, ya que según, Escobar-Pérez y Cuervo-Martínez (2008) y Torres-Malca et al. (2022) son opiniones bien informadas de personas que poseen experiencia y conocimientos en un campo determinado y que son reconocidas como especialistas. Estas personas son capaces de ofrecer datos, pruebas, opiniones y evaluaciones. El juicio de expertos es esencialmente la evaluación o el examen de un tema o cuestión por parte de quienes tienen conocimientos especializados en esa disciplina concreta. El proceso de juicio de expertos implica recopilar opiniones y evaluaciones de personas que tienen un profundo conocimiento en un tema en particular. Esto puede ser utilizado en una variedad de contextos, como en la toma de decisiones en la investigación científica, la resolución de problemas técnicos, la evaluación de riesgos, o en la toma de decisiones en áreas especializadas.

Los resultados de los expertos fueron los siguientes:

Expertos	Condición final
Vera Ramirez, Oscar John	Excelente
Flores Gutierrez, Vaneza	Muy buena
Flores Quispe, Eduardo Luis	Muy buena
Pinillos Gallardo, María del Rocio	Excelente
Oscó Mamani Eber Francisco	Excelente

### Confiabilidad

Según Carrasco (2006) es la fiabilidad a la capacidad de una herramienta de medición para producir resultados coherentes cuando se utiliza en varias ocasiones con los mismos sujetos en distintos periodos. En este caso, se utilizó el coeficiente de fiabilidad Alfa de Cronbach junto con el enfoque de consistencia interna para determinar la fiabilidad del cuestionario.

$$\alpha = \frac{k}{k-1} \left( 1 - \frac{\sum S_i^2}{S_T^2} \right)$$

Dónde:

K = Número de ítems que se consideran en el instrumento

SI<sup>2</sup> = La varianza del instrumento

ST<sup>2</sup> = La varianza de la suma de los ítems.

Según la herramienta estadística SPSS v 26 los resultados obtenidos fueron los siguientes:

Escala variable1: Productividad educativa

Estadísticas de fiabilidad	
Alfa de Cronbach	N de elementos
0.777	19

Escala variable 2: Algoritmos de Aprendizaje Automático

Estadísticas de fiabilidad	
Alfa de Cronbach	N de elementos
0.974	25

## Procesamiento y análisis de datos

El procesamiento y análisis de datos de la presente investigación, se realizó proporcionando una descripción general del proceso, teniendo en cuenta la herramienta SPSS v26:

- **Preparación de Datos:** Se realizó la recopilación de datos, limpieza de datos y Codificación de Variables, consistente en la asignación de códigos o etiquetas a las categorías de las variables para facilitar el análisis.
- **Exploración de Datos:** Consistió en realizar estadísticas descriptivas (utilización de medidas como media, mediana, desviación estándar, etc., para obtener una visión general de las variables) y visualización de Datos.
- **Análisis Estadístico:** Se realizó pruebas de Hipótesis, Análisis de Varianza y análisis de regresión, es decir la exploración de las relaciones entre las variables, especialmente la productividad educativa y los algoritmos de aprendizaje automático.
- **Implementación de Algoritmos de Aprendizaje Automático:** Selección de Modelos: Elección de algoritmos de aprendizaje automático adecuados para el análisis, teniendo en cuenta los objetivos de la investigación, entrenamiento del Modelo y validación del Modelo.
- **Uso de SPSS v26:** consistente en la iimportación de Datos: Cargar y organizar los datos en SPSS, análisis estadístico y generación de gráficos: Crear gráficos descriptivos y representativos de los resultados.

En cada etapa fue fundamental para la reflexión, validación de las suposiciones y considerando las limitaciones de los análisis realizados. La combinación de análisis estadísticos y técnicas de aprendizaje automático suministró una perspectiva integral sobre la relación entre la productividad educativa y los algoritmos de aprendizaje automático en la Escuela de Ingeniería de Sistemas e Informática de la UNAM en 2023.

## Ética en la investigación

El trabajo de investigación guardó los principios establecidos según lo que dispone la Universidad Alas Peruanas referente al citado y normas APA 7 y de propiedad intelectual.

Asimismo, Correia (2023) y Sivasubramaniam et al. (2021) destacan que la ética en la investigación y su práctica en el comportamiento de una sociedad civilizada es fundamental para descartar posibles malas conductas, sea en el método científico o con la forma en que se presentan los resultados al mundo.

Por tanto, el contenido del trabajo de investigación guarda los principios establecidos:

La información personal se mantuvo privada en todo momento mientras se utilizaba el dispositivo.



Esta investigación es totalmente original y escrita por mí; todas las citas, ya sean directas de la fuente o parafraseadas, han sido debidamente citadas.

Durante la elaboración de la presente investigación no se alteraron las respuestas ni se obtuvieron resultados sesgados de los instrumentos.

## Referencias

- Arias Gonzáles, J. L. (2021). *Diseño y metodología de la investigación*. Enfoques Consulting E.I.R.L. [https://apps.utel.edu.mx/recursos/files/r161r/w26022w/Arias\\_S2.pdf](https://apps.utel.edu.mx/recursos/files/r161r/w26022w/Arias_S2.pdf)
- Bernal Torres, C. A. (2016). *Metodología de la investigación: Administración, economía, humanidades y ciencias sociales* (4ª ed.). Pearson Educación.  
<https://drive.google.com/file/d/1-3wqx7vGGCn6O4FxMPkzKw15E4tByYXX/view?pli=1>
- Carrasco Díaz, S. (2006). *Metodología de la investigación científica: pautas metodológicas para diseñar y elaborar el proyecto de investigación*. San Marcos.  
<https://drive.google.com/file/d/1GTWMTyAZDmzEohJbUKSxsR-QJWsYugBV/view>
- Correia, M. I. T. D. (2023). Ethics in research. *Clinical Nutrition Open Science*, 47, 121–130.  
<https://doi.org/10.1016/j.nutos.2022.12.010>
- Escobar-Pérez, J. y Cuervo-Martínez, Á. (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. *Avances En Medición*, 27–36.  
[https://www.humanas.unal.edu.co/lab\\_psicometria/application/files/9416/0463/3548/Vol\\_6.\\_Articulo3\\_Juicio\\_de\\_expertos\\_27-36.pdf](https://www.humanas.unal.edu.co/lab_psicometria/application/files/9416/0463/3548/Vol_6._Articulo3_Juicio_de_expertos_27-36.pdf)
- Hernández Sampieri, R. y Fernández Collado, C. y Baptista Lucio, M. del P. (2014). *Metodología de la investigación*. Mc Graw Hill.
- Jin, S. y Fang, G. y Cheung, K. C. y Sit, P. S. (2022). Factors associated with academic resilience in disadvantaged students: An analysis based on the PISA 2015 B-S-J-G (China) sample. *Frontiers in Psychology*, 13.  
<https://doi.org/10.3389/fpsyg.2022.846466>
- Maina, J. J. y Zakari, A. T. y Alkali, I. A. y Salisu, R. A. (2021). Academic success predictors for architecture students at Kano University of Science and Technology, Wudil, Kano State, Nigeria. *Bayero Journal of Pure and Applied Sciences*, 13(2), 125–133.  
<https://doi.org/10.4314/bajopas.v13i2.17>
- Maletta, H. (2000). *Epistemología aplicada: notas de clase*. Universidad del Salvador.  
<http://blog.pucp.edu.pe/blog/maestriaeducacion/wp-content/uploads/sites/184/2012/01/epistemologiaplicada-Maletta.pdf>

- Palanichamy, K. (2019). Integrative Omic Analysis of Neuroblastoma. In *Computational Epigenetics and Diseases* (pp. 311–326). Elsevier. <https://doi.org/10.1016/B978-0-12-814513-5.00019-2>
- Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, 2(3), 160. <https://doi.org/10.1007/s42979-021-00592-x>
- Sivasubramaniam, S. y Dlabolová, D. H. y Kralikova, V. y Khan, Z. R. (2021). Assisting you to advance with ethics in research: an introduction to ethical governance and application procedures. *International Journal for Educational Integrity*, 17(1), 14. <https://doi.org/10.1007/s40979-021-00078-6>
- Tamayo y Tamayo, M. (2004). *El proceso de la investigación científica*. Limusa. [https://www.academia.edu/29308889/Tamayo\\_Mario\\_El\\_Proceso\\_De\\_La\\_Investigacion\\_Cientifica\\_pdf](https://www.academia.edu/29308889/Tamayo_Mario_El_Proceso_De_La_Investigacion_Cientifica_pdf)
- Torres-Malca, J. R. y Vera-Ponce, V. J. y Zuzunaga-Montoya, F. E. y Talavera, J. E. y De La Cruz-Vargas, J. A. (2022). Content validity by expert judgment of an instrument to measure knowledge, attitudes and practices about salt consumption in the peruvian population. *Revista de La Facultad de Medicina Humana*, 22(2), 273–279. <https://doi.org/10.25176/RFMH.v22i2.4768>

# CAPÍTULO 5

## Resultados

### Results

DOI: <https://doi.org/10.71112/23nxem35>

#### Resumen

Este capítulo presenta los resultados y análisis de la investigación sobre la relación entre la productividad educativa y los algoritmos de aprendizaje automático. El análisis descriptivo detalla las percepciones de los 90 estudiantes encuestados sobre dimensiones como aptitud, instrucción, medio ambiente y aprendizaje. El análisis inferencial, mediante la prueba Rho de Spearman, revela una correlación positiva muy alta (0.940) entre ambas variables, confirmando todas las hipótesis específicas. Se concluye que los algoritmos (Decision Trees, Random Forest, Naive Bayes) son predictores efectivos de la productividad educativa, validando estudios previos.

**Palabras clave:** Resultados, Correlación, Rho de Spearman, Análisis Inferencial, Productividad Educativa, Aprendizaje Automático.

#### Abstract

This chapter presents the results and analysis of the research on the relationship between educational productivity and machine learning algorithms. The descriptive analysis details the perceptions of the 90 surveyed students regarding dimensions such as aptitude, instruction, environment, and learning. The inferential analysis, using Spearman's Rho test, reveals a very high positive correlation (0.940) between both variables, confirming all specific hypotheses. It is concluded that the algorithms (Decision Trees, Random Forest, Naive Bayes) are effective predictors of educational productivity, validating previous studies.

**Keywords:** Results, Correlation, Spearman's Rho, Inferential Analysis, Educational Productivity, Machine Learning.

#### Análisis descriptivo

Luego de la aplicación de la herramienta de recolección de datos, se realizó el procesamiento necesario para analizar los datos. En este punto, se analizaron internamente todos los indicadores suministrados por la EPISI de la (UNAM).

## Descripción de encuestados

**Tabla 2**

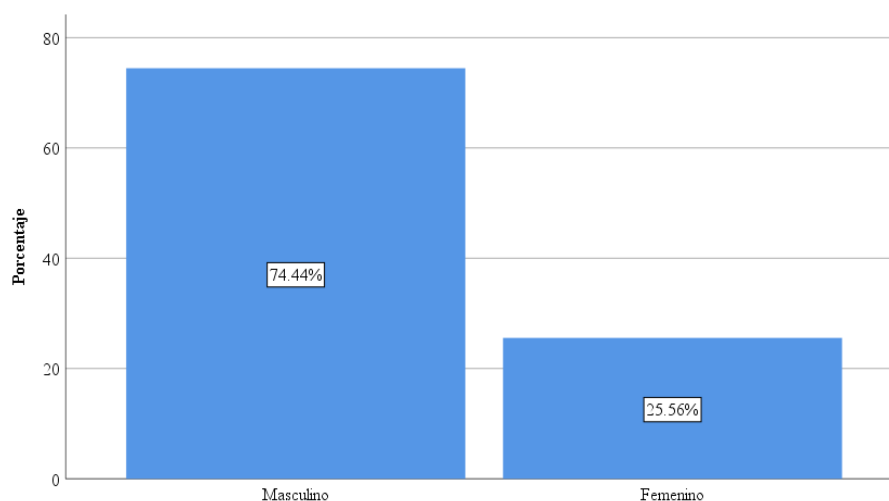
*Totalidad de encuestados por genero*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Masculino	67	74.4	74.4	74.4
	Femenino	23	25.6	25.6	100.0
	Total	90	100.0	100.0	

*Nota:* En la tabla podemos observar la distribución de encuestados por frecuencia de género, de ellos 67 son Masculinos y 23 Femeninos que representan el 74.4% y 23% respectivamente en la presente investigación.

**Figura 6**

Encuestados según genero



*Nota:* La figura describe porcentualmente la muestra de representativa de los alumnos de la EPISI de la UNAM según genero considerado en la presente investigación.

## Variable 1: productividad educativa

**Tabla 3**

*Pregunta 1. Utilizo los recursos académicos disponibles para mis conocimientos*

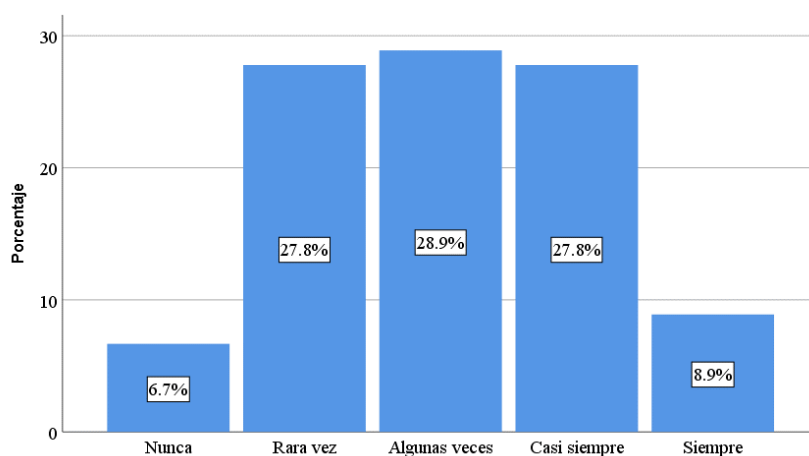
	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Nunca	6	6,7	6,7	6,7

Válido	Rara Vez	25	27,8	27,8	34,4
	Algunas Veces	26	28,9	28,9	63,3
	Casi Siempre	25	27,8	27,8	91,1
	Siempre	8	8,9	8,9	100,0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 7**

*Pregunta 1. Utilizo los recursos académicos disponibles para mis conocimientos*



*Nota:* Distribución porcentual de la pregunta 1.

Interpretación:

De acuerdo a la pregunta 1, acerca del uso de los recursos académicos disponibles para sus conocimientos podemos mencionar que 28.9% (26 encuestados) mencionan algunas veces, 27.8% (25 encuestados) mencionan casi siempre y rara vez, 8.9% (8 encuestado) mencionan siempre y 6.7% (6 encuestados) mencionan no utilizan.

**Tabla 4**

*Pregunta 2. Planifico mis estudios académicos en la universidad*

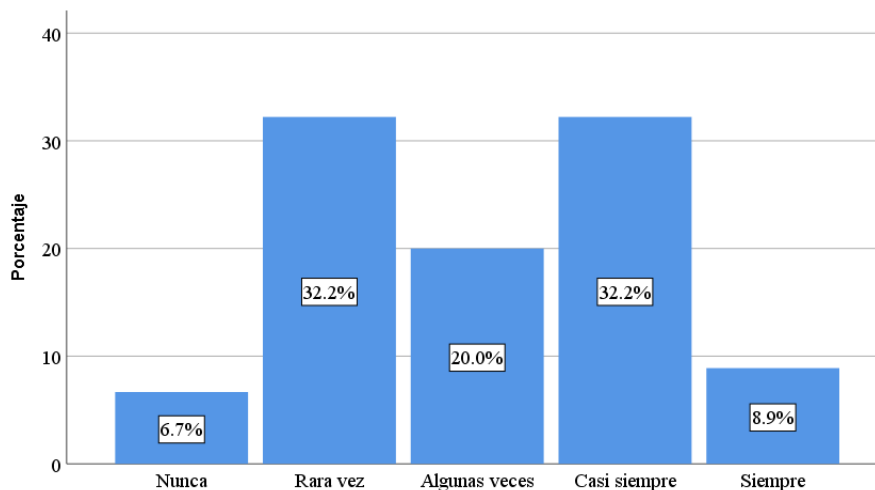
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Nunca	6	6,7	6,7	6,7
	Rara Vez	29	32,2	32,2	38,9
	Algunas Veces	18	20,0	20,0	58,9

Casi Siempre	29	32,2	32,2	91,1
Siempre	8	8,9	8,9	100,0
Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 8**

*Pregunta 2. Planifico mis estudios académicos en la universidad*



*Nota:* Distribución porcentual de la pregunta 2.

Interpretación:

De acuerdo a la pregunta 2, acerca de la planificación de estudios en la universidad podemos mencionar que 32.2% (29 encuestados) mencionan rara vez y casi siempre, el 20% (18 encuestados) mencionan algunas veces, el 8.9% (8 encuestados) mencionan siempre y 6.7% (6 encuestados) mencionan nunca.

**Tabla 5**

*Pregunta 3: Los docentes felicitan mis esfuerzos académicos*

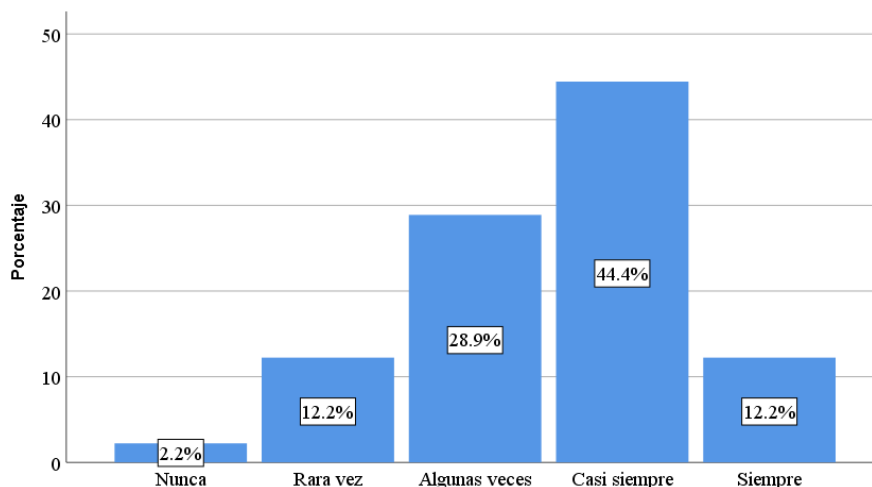
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Nunca	2	2,2	2,2	2,2
	Rara Vez	11	12,2	12,2	14,4
	Algunas Veces	26	28,9	28,9	43,3
	Casi Siempre	40	44,4	44,4	87,8

	Siempre	11	12,2	12,2	100,0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

### Figura 9

*Pregunta 3: Los docentes felicitan mis esfuerzos académicos*



*Nota:* Distribución porcentual de la pregunta 3.

Interpretación:

De acuerdo a la pregunta 3, referente al elogio establecido por los docentes al esfuerzo académico podemos mencionar que 44.4% (40 encuestados) mencionan casi siempre, el 28.9% (26 encuestados) mencionan algunas veces, mientras que el 12.2% (11 encuestados) mencionan rara vez y siempre y 2.2% (2 encuestados) mencionan nunca.

### Tabla 6

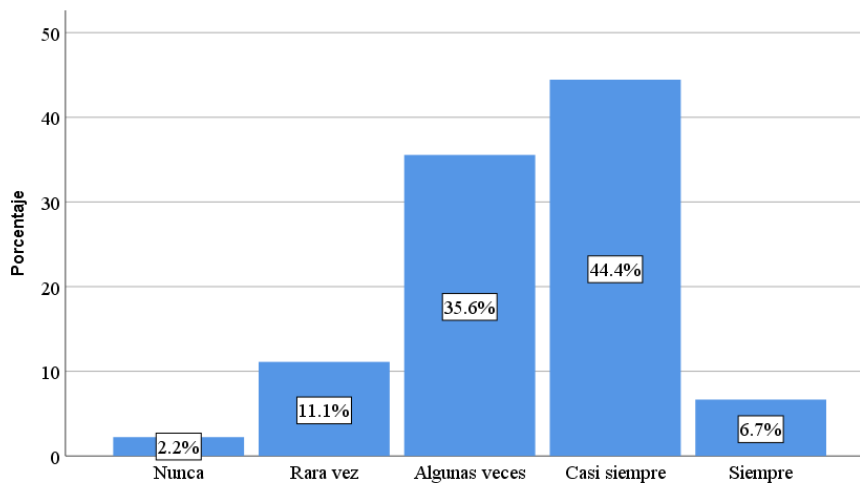
*Pregunta 4: Comprendo las sesiones de aprendizaje*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Nunca	2	2,2	2,2	2,2
	Rara Vez	10	11,1	11,1	13,3
	Algunas Veces	32	35,6	35,6	48,9
	Casi Siempre	40	44,4	44,4	93,3
	Siempre	6	6,7	6,7	100,0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 10**

*Pregunta 4: Comprendo las sesiones de aprendizaje*



*Nota:* Distribución porcentual de la pregunta 4.

Interpretación:

De acuerdo a la pregunta 4, acerca de la comprensión de las sesiones de aprendizaje mencionan que el 44.4% (40 encuestados) mencionan casi siempre, 35.6% (32 encuestados) mencionan algunas veces, 11.1% (10 encuestado) mencionan rara vez, 6.7% (6 encuestado) siempre y 2.2% (2 encuestados) mencionan nunca.

**Tabla 7**

*Pregunta 5: Planifico mis tareas académicas*

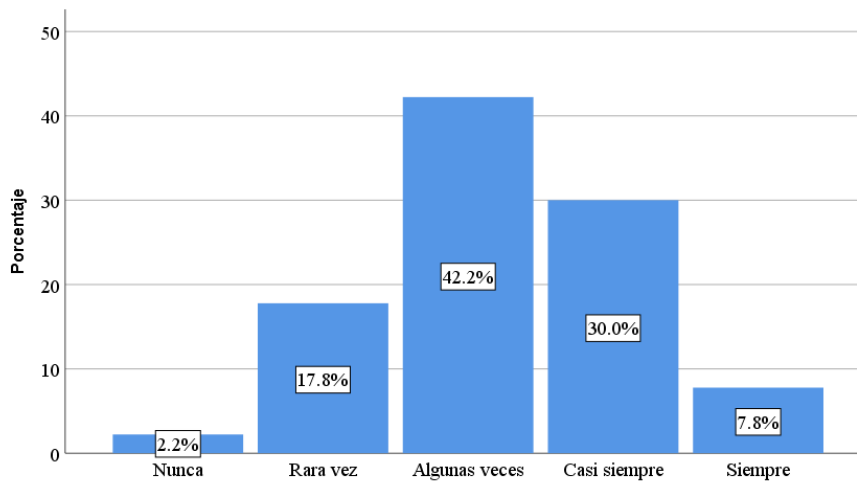
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Nunca	2	2,2	2,2	2,2
	Rara Vez	16	17,8	17,8	20,0
	Algunas Veces	38	42,2	42,2	62,2
	Casi Siempre	27	30,0	30,0	92,2
	Siempre	7	7,8	7,8	100,0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.



**Figura 11**

*Pregunta 5: Planifico mis trabajos académicos*



*Nota:* Distribución porcentual de la pregunta 5.

Interpretación:

De acuerdo a la pregunta 5, acerca de la planificación de las tareas académicas mencionan que el 42.2% (38 encuestados) mencionan algunas veces, el 30.0% (27 encuestados) mencionan casi siempre, 17.8% (16 encuestado) mencionan rara vez, 7.8% (7 encuestado) siempre y 2.2% (2 encuestados) mencionan nunca.

**Tabla 8**

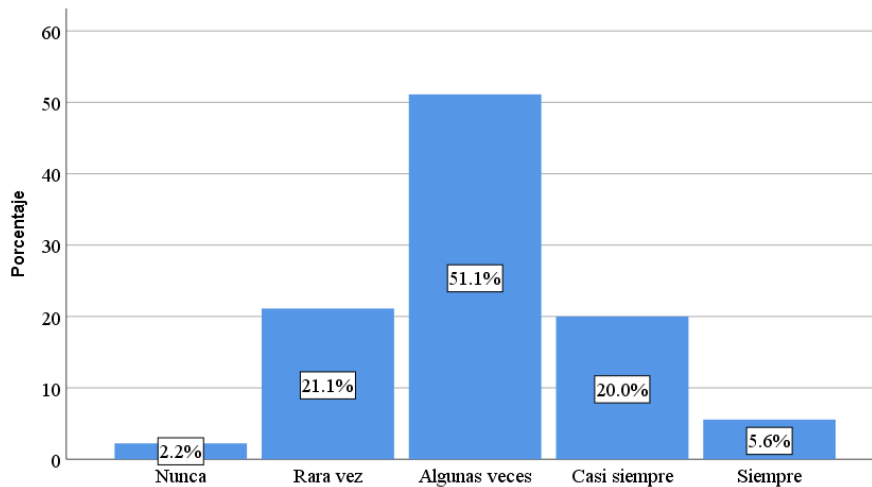
*Pregunta 6: Utilizo la biblioteca para ampliar mis conocimientos*

			Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
o	Válid	Nunca	2	2,2	2,2	2,2
		Rara Vez	19	21,1	21,1	23,3
		Algunas Veces	46	51,1	51,1	74,4
		Casi Siempre	18	20,0	20,0	94,4
		Siempre	5	5,6	5,6	100,0
		Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 12**

*Pregunta 6: Utilizo la biblioteca para ampliar mis conocimientos*



*Nota:* Distribución porcentual de la pregunta 6.

Interpretación:

De acuerdo a la pregunta 6, acerca del uso de la biblioteca para ampliar los conocimientos mencionan que el 51.1% (46 encuestados) mencionan algunas veces, el 21.1% (19 encuestados) mencionan rara vez, el 20% (18 encuestados) mencionan casi siempre, el 5.6% (5 encuestados) siempre y el 2.2% (2 encuestados) mencionan nunca.

**Tabla 9**

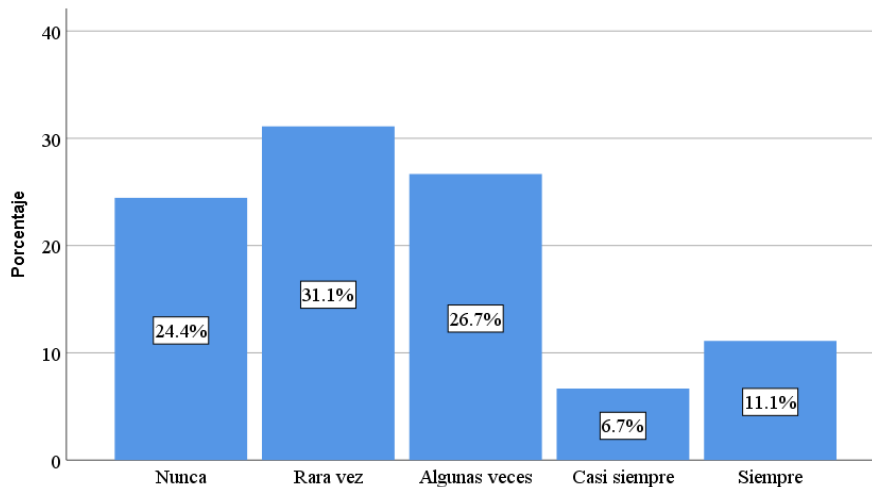
*Pregunta 7: Mis padres apoyan en mis conocimientos académicos*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Nunca	22	24,4	24,4	24,4
	Rara Vez	28	31,1	31,1	55,6
	Algunas Veces	24	26,7	26,7	82,2
	Casi Siempre	6	6,7	6,7	88,9
	Siempre	10	11,1	11,1	100,0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 13**

*Pregunta 7: Mis padres apoyan en mis conocimientos académicos*



*Nota:* Distribución porcentual de la pregunta 7.

Interpretación:

De acuerdo a la pregunta 7, acerca del apoyo de los padres en el conocimiento académico mencionan que el 31.1% (28 encuestados) mencionan rara vez, luego el 26.7% (24 encuestados) mencionan algunas veces, el 24.4% (22 encuestados) mencionan nunca, el 11.1% (10 encuestado) siempre y el 6.7% (6 encuestados) mencionan casi siempre.

**Tabla 10**

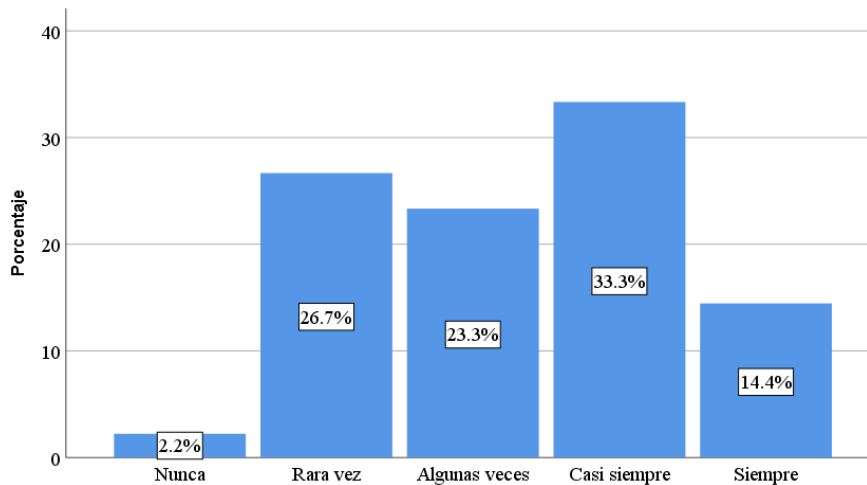
*Pregunta 8: Utilizo equipos y herramientas académicas de casa para estudiar*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Nunca	2	2,2	2,2	2,2
	Rara Vez	24	26,7	26,7	28,9
	Algunas Veces	21	23,3	23,3	52,2
	Casi Siempre	30	33,3	33,3	85,6
	Siempre	13	14,4	14,4	100,0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 14**

*Pregunta 8: Utilizo equipos y herramientas académicas de casa para estudiar*



*Nota:* Distribución porcentual de la pregunta 8.

Interpretación:

De acuerdo a la pregunta 8, acerca del uso de equipos y herramientas académicas para estudiar mencionan que el 33.3% (30 encuestados) mencionan casi siempre, mientras que el 26.7% (24 encuestados) mencionan rara vez, el 23.3% (21 encuestados) mencionan algunas veces, el 14.4% (13 encuestados) mencionan siempre y el 2.2% (2 encuestados) mencionan nunca.

**Tabla 11**

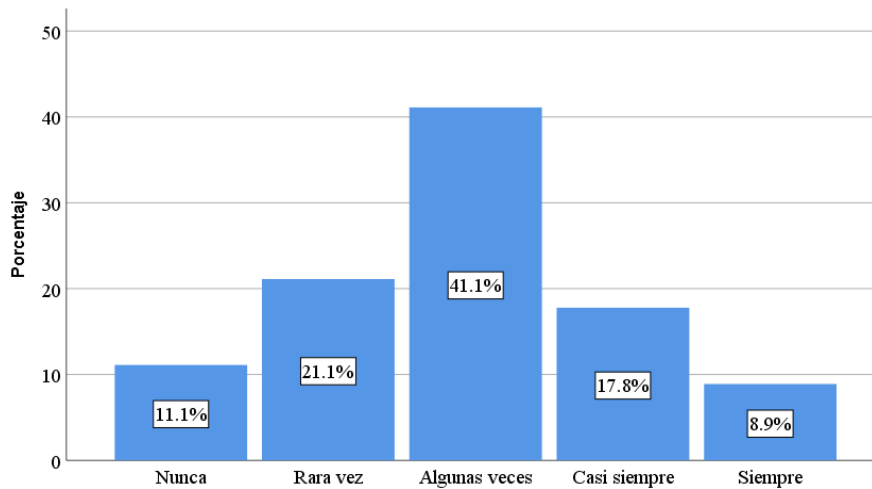
*Pregunta 9: Recibo apoyo financiero familiar para mis estudios*

			Porcentaje	Porcentaje
			válido	acumulado
Válido	Nunca	10	11,1	11,1
	Rara Vez	19	21,1	32,2
	Algunas Veces	37	41,1	73,3
	Casi Siempre	16	17,8	91,1
	Siempre	8	8,9	100,0
	Total	90	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 15**

*Pregunta 9: Recibo apoyo financiero familiar para mis estudios*



*Nota:* Distribución porcentual de la pregunta 9.

Interpretación:

De acuerdo a la pregunta 9, acerca del apoyo financiero para sus estudios mencionan que el 41.1% (37 encuestados) mencionan algunas veces, mientras que el 21.1% (19 encuestados) mencionan rara vez, el 17.8% (16 encuestados) mencionan casi siempre, el 11.1% (10 encuestados) mencionan nunca y el 8.9% (8 encuestados) mencionan siempre.

**Tabla 12**

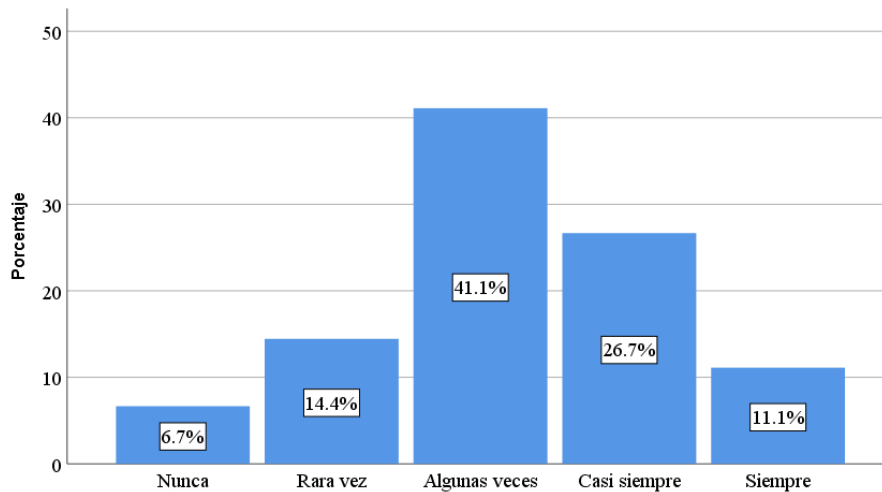
*Pregunta 10: Realizo desenvolvimiento social con mis vecinos*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Nunca	6	6,7	6,7	6,7
	Rara Vez	13	14,4	14,4	21,1
	Algunas Veces	37	41,1	41,1	62,2
	Casi Siempre	24	26,7	26,7	88,9
	Siempre	10	11,1	11,1	100,0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 16**

*Pregunta 10: Realizo desenvolvimiento social con mis vecinos*



*Nota:* Distribución porcentual de la pregunta 10.

Interpretación:

De acuerdo a la pregunta 10, acerca del desenvolvimiento social con los vecinos del encuestado manifiestan que el 41.1% (37 encuestados) mencionan algunas veces, mientras que el 26.7% (24 encuestados) mencionan casi siempre, el 14.4% (13 encuestados) mencionan rara vez, el 11.1% (10 encuestados) mencionan siempre y el 6.7% (6 encuestados) mencionan nunca.

**Tabla 13**

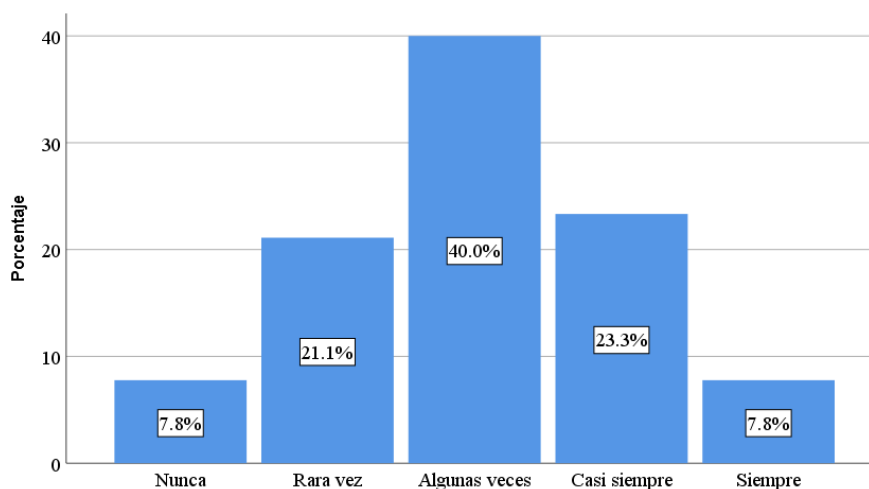
*Pregunta 11: Participo en actividades deportivas en mi barrio*

			Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Nunca		7	7,8	7,8	7,8
	Rara Vez		19	21,1	21,1	28,9
	Algunas Veces		36	40,0	40,0	68,9
	Casi Siempre		21	23,3	23,3	92,2
	Siempre		7	7,8	7,8	100,0
	Total		90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 17**

*Pregunta 11: Participo en actividades deportivas en mi barrio*



*Nota:* Distribución porcentual de la pregunta 11.

Interpretación:

De acuerdo a la pregunta 11, acerca de la participación en actividades deportivas en sus barrios manifiestan que el 40% (36 encuestados) mencionan algunas veces, mientras que el 23.3% (21 encuestados) mencionan casi siempre, el 19% (19 encuestados) mencionan rar vez, el 7.8% (7 encuestados) mencionan siempre y nunca.

**Tabla 14**

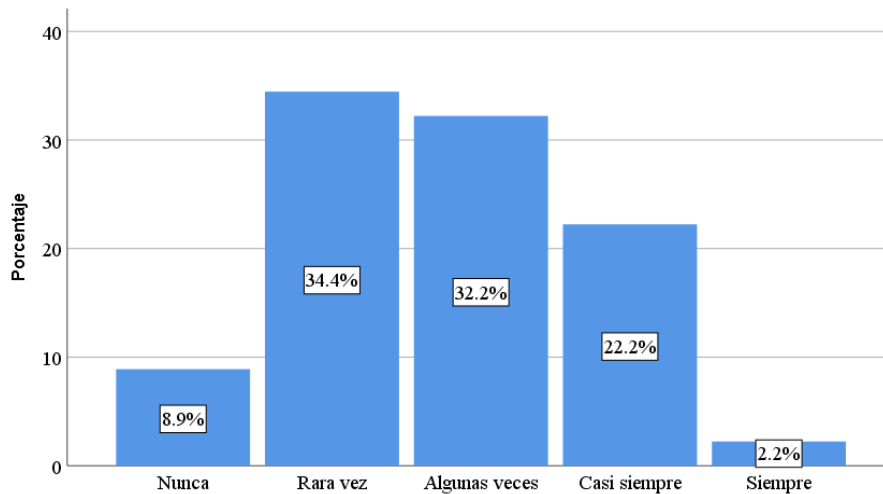
*Pregunta 12: Participo en actividades comunitarias con los conocimientos adquiridos*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Nunca	8	8,9	8,9	8,9
	Rara Vez	31	34,4	34,4	43,3
	Algunas Veces	29	32,2	32,2	75,6
	Casi Siempre	20	22,2	22,2	97,8
	Siempre	2	2,2	2,2	100,0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 18**

*Pregunta 12: Participo en actividades comunitarias con los conocimientos adquiridos*



*Nota:* Distribución porcentual de la pregunta 12.

Interpretación:

De acuerdo a la pregunta 12, acerca de la participación en actividades comunitarias con los conocimientos adquiridos manifiestan que el 34.4% (31 encuestados) mencionan rara vez, mientras que el 32.2% (29 encuestados) mencionan algunas veces, el 22.2% (20 encuestados) mencionan casi siempre, el 8.9% (8 encuestados) mencionan nunca y el 2.2% (2 encuestados) mencionaron siempre.

**Tabla 15**

*Pregunta 13: Los medios de comunicación apoyan el desarrollo de mis conocimientos*

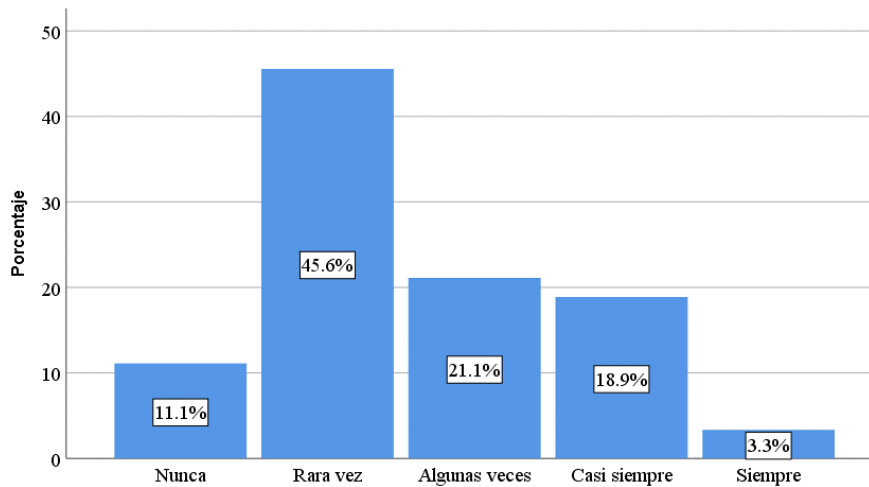
		Frecuencia	Porcentaje válido	Porcentaje	Porcentaje acumulado
Válido	Nunca	10	11,1	11,1	11,1
	Rara Vez	41	45,6	45,6	56,7
	Algunas Veces	19	21,1	21,1	77,8
	Casi Siempre	17	18,9	18,9	96,7
	Siempre	3	3,3	3,3	100,0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.



**Figura 19**

*Pregunta 13: Los medios de comunicación apoyan el desarrollo de mis conocimientos*



*Nota:* Distribución porcentual de la pregunta 13.

Interpretación:

De acuerdo a la pregunta 13, acerca de los aportes de medios de comunicación en el desarrollo del conocimiento manifiestan que el 45.6% (41 encuestados) mencionan rara vez, mientras que el 21.1% (19 encuestados) mencionan algunas veces, el 18.9% (17 encuestados) mencionan casi siempre, el 11.1% (10 encuestados) mencionan nunca y el 3.3% (3 encuestados) mencionaron siempre.

**Tabla 16**

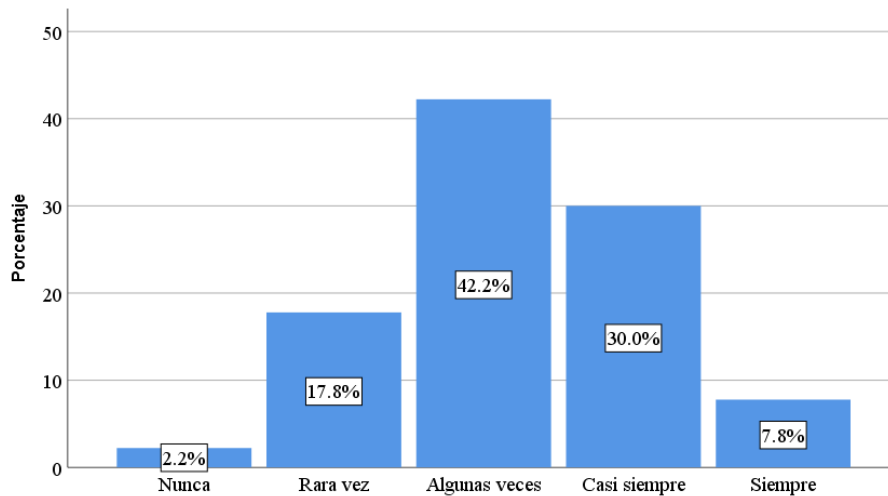
*Pregunta 14: Planifico mis actividades académicas constantemente*

			Porcentaje	Porcentaje
			válido	acumulado
Válido	Nunca	2	2,2	2,2
	Rara Vez	16	17,8	20,0
	Algunas Veces	38	42,2	62,2
	Casi Siempre	27	30,0	92,2
	Siempre	7	7,8	100,0
	Total	90	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 20**

*Pregunta 14: Planifico mis actividades académicas constantemente*



*Nota:* Distribución porcentual de la pregunta 14.

Interpretación:

De acuerdo a la pregunta 14, acerca de planificar los estudios constantemente manifiestan que el 42.2.6% (38 encuestados) mencionan algunas veces, mientras que el 30% (27 encuestados) mencionan casi siempre, el 17.8% (16 encuestados) mencionan rara vez, el 7.8% (7 encuestados) mencionan siempre y el 2.2% (2 encuestados) mencionaron nunca.

**Tabla 17**

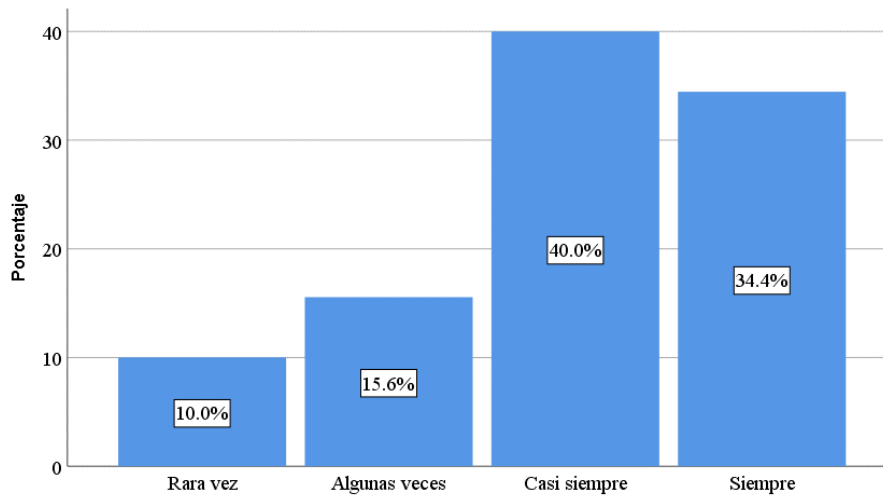
*Pregunta 15: Planifico el uso de TICs para mis estudios e investigaciones*

			Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Rara Vez		9	10,0	10,0	10,0
	Algunas Vezes		14	15,6	15,6	25,6
	Casi Siempre		36	40,0	40,0	65,6
	Siempre		31	34,4	34,4	100,0
	Total		90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 21**

*Pregunta 15: Planifico el uso de TICs para mis estudios e investigaciones*



*Nota:* Distribución porcentual de la pregunta 15.

Interpretación:

De acuerdo a la pregunta 15, acerca de planificar el uso de TICs para estudios e investigaciones manifiestan que el 40% (36 encuestados) mencionan casi siempre, mientras que el 34.4% (31 encuestados) mencionan siempre, el 15.6% (14 encuestados) mencionan algunas veces y el 10% (9 encuestados) mencionan rara vez.

**Tabla 18**

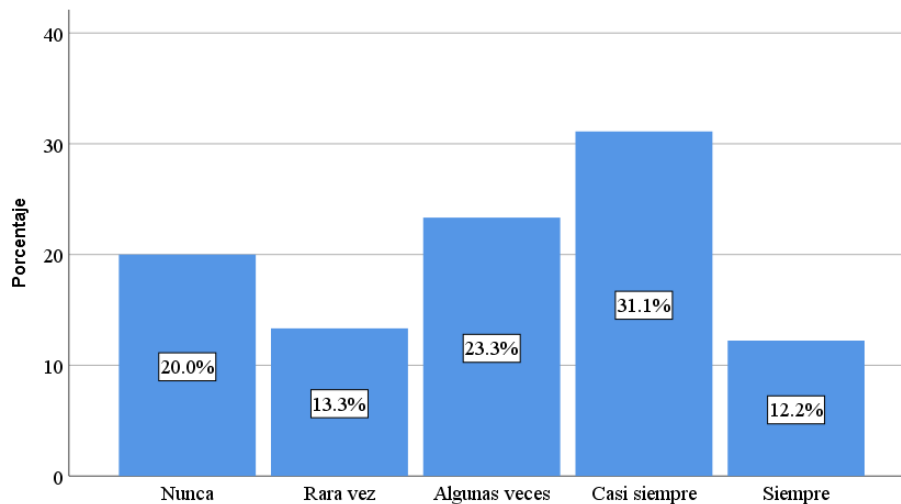
*Pregunta 16: Planifico realizar proyectos de emprendimiento académico*

			Porcentaje	Porcentaje
			válido	acumulado
Frecuencia			Porcentaje	
Válido	Nunca	18	20,0	20,0
	Rara Vez	12	13,3	33,3
	Algunas Veces	21	23,3	56,7
	Casi Siempre	28	31,1	87,8
	Siempre	11	12,2	100,0
	Total	90	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 22**

*Pregunta 16: Planifico realizar proyectos de emprendimiento académico*



*Nota:* Distribución porcentual de la pregunta 16.

Interpretación:

De acuerdo a la pregunta 16, acerca de planificar proyectos de emprendimiento académico manifiestan que el 31.1% (28 encuestados) mencionan casi siempre, mientras que el 23.3% (21 encuestados) mencionan algunas veces, el 20% (18 encuestados) mencionan nunca, el 13.3% (12 encuestados) mencionan rara vez y el 12.2% (11 encuestados) mencionan siempre.

**Tabla 19**

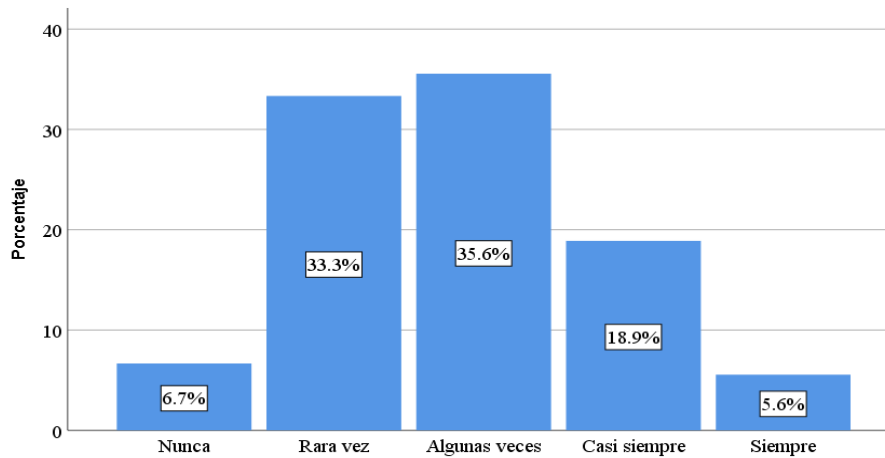
*Pregunta 17: Existe socialización académica con los docentes*

			Porcentaje	Porcentaje
			válido	acumulado
Válido	Nunca	6	6,7	6,7
	Rara Vez	30	33,3	40,0
	Algunas Veces	32	35,6	75,6
	Casi Siempre	17	18,9	94,4
	Siempre	5	5,6	100,0
	Total	90	100,0	100,0

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 23**

*Pregunta 17: Existe socialización académica con los docentes*



*Nota:* Distribución porcentual de la pregunta 17.

Interpretación:

De acuerdo a la pregunta 17, acerca de la socialización académica con los docentes manifiestan que el 35.6% (32 encuestados) mencionan algunas veces, mientras que el 33.3% (30 encuestados) mencionan rara vez, el 18.9% (17 encuestados) mencionan casi siempre, el 6.7% (6 encuestados) mencionan nunca y el 5.6% (5 encuestados) mencionan siempre.

**Tabla 20**

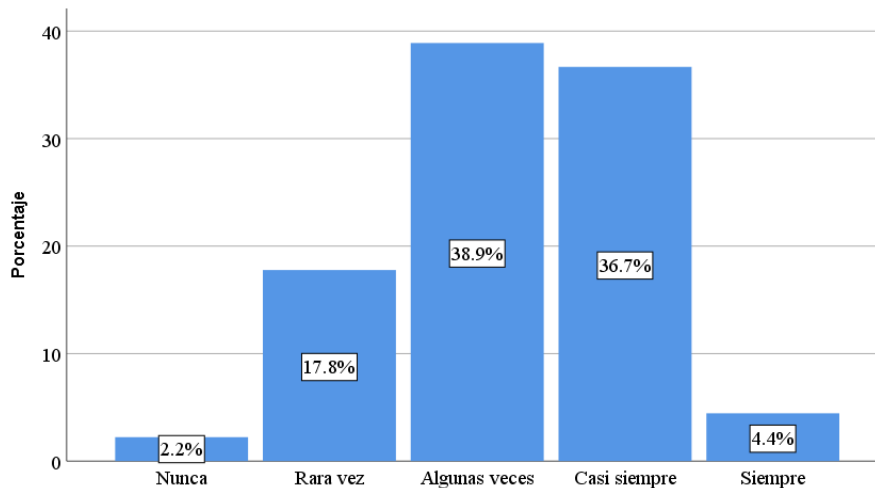
*Pregunta 18: Existe satisfacción académica en mi experiencia universitaria*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Nunca	2	2,2	2,2	2,2
	Rara Vez	16	17,8	17,8	20,0
	Algunas Veces	35	38,9	38,9	58,9
	Casi Siempre	33	36,7	36,7	95,6
	Siempre	4	4,4	4,4	100,0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 24**

*Pregunta 18: Existe satisfacción académica en mi experiencia universitaria*



*Nota:* Distribución porcentual de la pregunta 18.

#### Interpretación:

De acuerdo a la pregunta 18, acerca de la satisfacción académica como experiencia universitaria en general manifiestan que el 38.9% (35 encuestados) mencionan algunas veces, mientras que el 36.7% (33 encuestados) mencionan casi siempre, el 17.8% (16 encuestados) mencionan rara vez, el 4.4% (4 encuestados) mencionan siempre y el 2.2% (2 encuestados) mencionan nunca.

**Tabla 21**

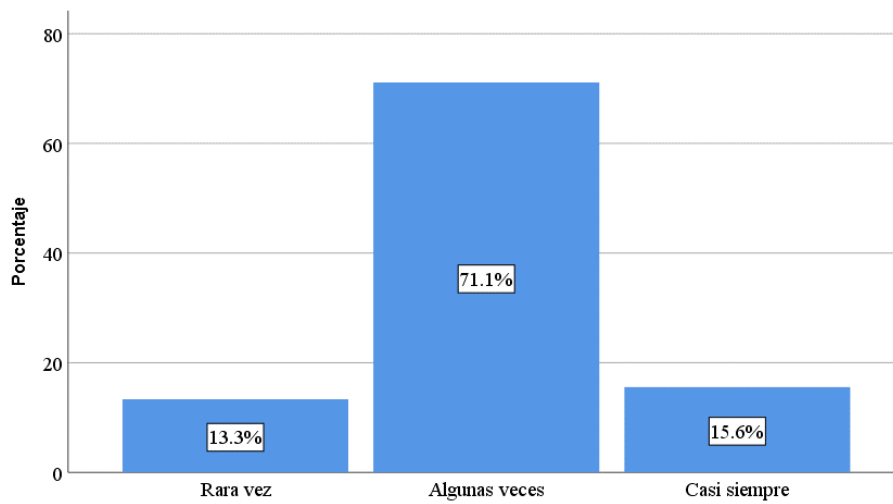
*Pregunta 19: El nivel de logro de aprendizaje es satisfactorio*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Insuficiente	12	13.3	13.3	13.3
	Aprobado	64	71.1	71.1	84.4
	Bueno	14	15.6	15.6	100.0
	Total	90	100.0	100.0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 25**

*Pregunta 19: El nivel de logro de aprendizaje es satisfactorio*



*Nota:* Distribución porcentual de la pregunta 19.

Interpretación:

Conforme a la pregunta 19, acerca del logro de aprendizaje obtenido, en general manifiestan que el 13.3% (12 encuestados) mencionan rara vez, mientras que el 71.1% (64 encuestados) mencionan algunas veces y el 15.6% (14 encuestados) mencionan que el logro fue casi siempre.

## **Variable 2: algoritmos de aprendizaje automático**

**Tabla 22**

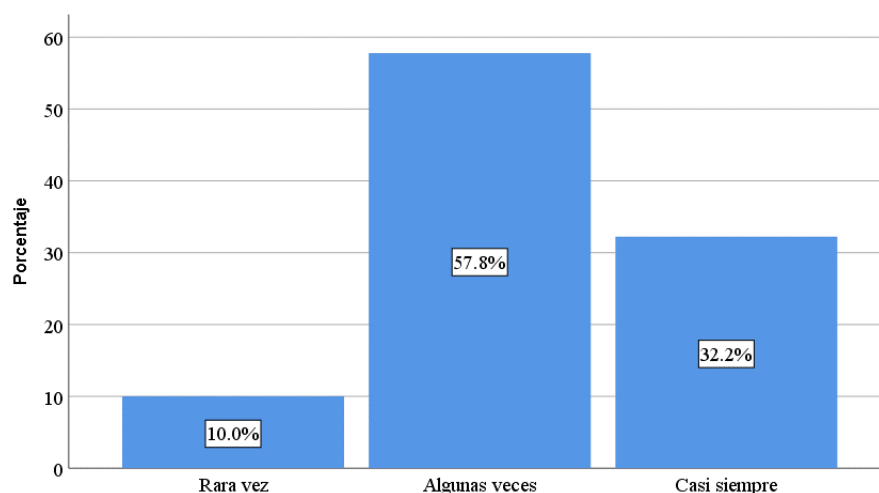
*Po1: Aprendizaje dimensión actitud del algoritmo arboles de decisión*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Valido	Rara vez	9	10.0	10.0	10.0
	Algunas veces	52	57.8	57.8	67.8
	Casi siempre	29	32.2	32.2	100.0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 26**

*Po1: Aprendizaje dimensión actitud del algoritmo arboles de decisión*



*Nota:* Distribución porcentual de la pregunta 01.

Interpretación:

De acuerdo al criterio 01, referente al aprendizaje en la dimensión actitud del algoritmo arboles de decisión establece que el 57.8% (52 encuestados) mencionan algunas veces, mientras que el 32.2% (29 encuestados) mencionan casi siempre y el 10% (9 encuestados) mencionan rara vez.

**Tabla 23**

*Po2: Aprendizaje dimensión medio ambiente del algoritmo arboles de decisión*

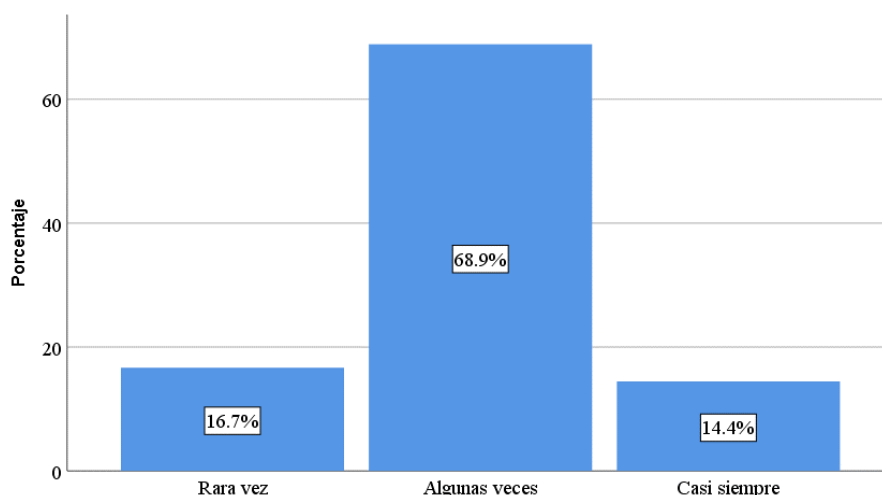
		Frecuencia	Porcentaje	Porcentaje e válido	Porcentaje acumulado
Válido	Rara vez	15	16.7	16.7	16.7
	Algunas veces	62	68.9	68.9	85.6
	Casi siempre	13	14.4	14.4	100.0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.



**Figura 27**

*Po2: Aprendizaje dimensión medio ambiente del algoritmo arboles de decisión*



*Nota:* Distribución porcentual de la pregunta 02.

Interpretación:

De acuerdo al criterio 02, referente al aprendizaje en la dimensión medio ambiente del algoritmo arboles de decisión establece que el 68.9% (62 encuestados) mencionan algunas veces, mientras que el 16.7% (15 encuestados) mencionan rara vez y el 14.4% (13 encuestados) mencionan casi siempre.

**Tabla 24**

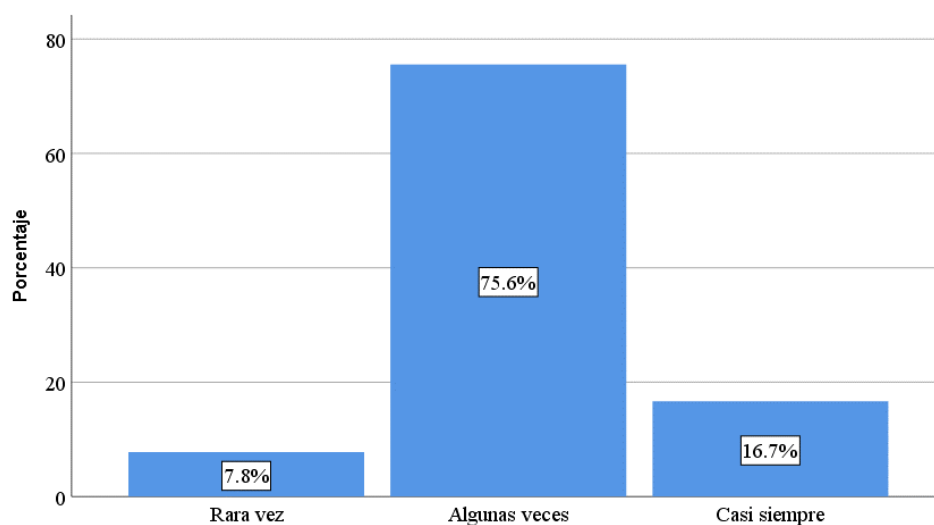
*Po3: Aprendizaje dimensión instrucción del algoritmo arboles de decisión*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Rara vez	7	7.8	7.8	7.8
	Algunas veces	68	75.6	75.6	83.3
	Casi siempre	15	16.7	16.7	100.0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 28**

*Po3: Aprendizaje dimensión instrucción del algoritmo arboles de decisión*



*Nota:* Distribución porcentual de la pregunta 03.

Interpretación:

De acuerdo al criterio 03, referente al aprendizaje en la dimensión instrucción del algoritmo arboles de decisión establece que el 75.6% (68 encuestados) mencionan algunas veces, mientras que el 16.7% (15 encuestados) mencionan casi siempre y el 7.8% (7 encuestados) mencionan rara vez.

**Tabla 25**

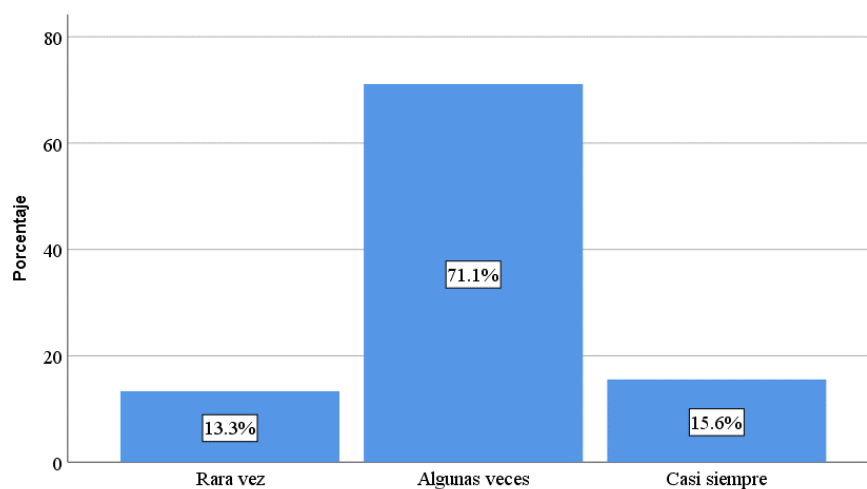
*Po4: Aprendizaje Logro del algoritmo arboles de decisión*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Rara vez	12	13.3	13.3	13.3
	Algunas veces	64	71.1	71.1	84.4
	Casi siempre	14	15.6	15.6	100.0
	Total	90	100.0	100.0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 29**

*Po4: Aprendizaje Logro del algoritmo arboles de decisión*



*Nota:* Distribución porcentual de la pregunta 04.

Interpretación:

De acuerdo al criterio 04, referente al aprendizaje en la dimensión logro del algoritmo arboles de decisión establecen que el 13.3% (12 encuestados) mencionan rara vez, mientras que el 71.1% (64 encuestados) mencionan algunas veces y el 15.6% (14 encuestados) casi siempre.

**Tabla 26**

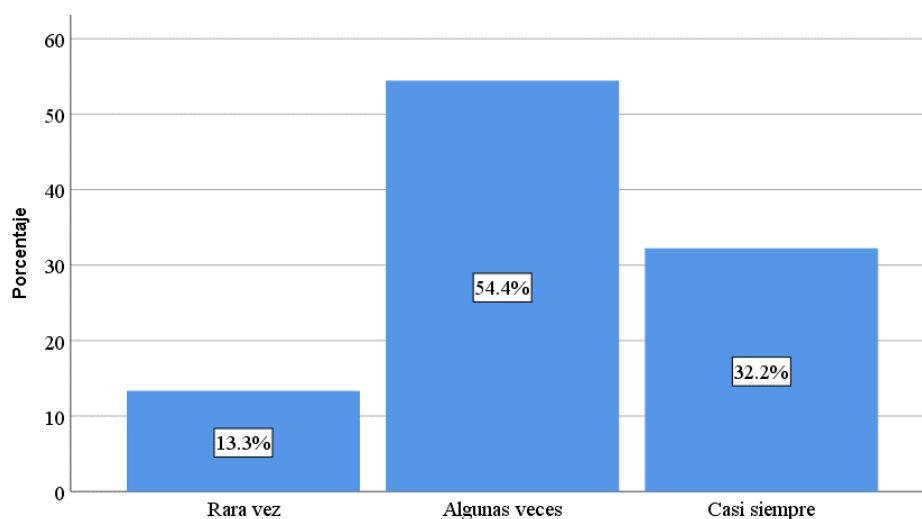
*Po5: Asertividad dimensión aptitud del algoritmo arboles de decisión*

			Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Rara vez		12	13.3	13.3	13.3
	Algunas veces		49	54.4	54.4	67.8
	Casi siempre		29	32.2	32.2	100.0
	Total		90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 30**

*P05: Asertividad dimensión aptitud del algoritmo arboles de decisión*



*Nota:* Distribución porcentual de la pregunta 05.

Interpretación:

De acuerdo al criterio 05, de asertividad en la dimensión aptitud del algoritmo arboles de decisión establece que el 54.4% (49 encuestados) mencionan algunas veces, mientras que el 32.2% (29 encuestados) mencionan casi siempre y el 13.3% (12 encuestados) mencionan rara vez.

**Tabla 27**

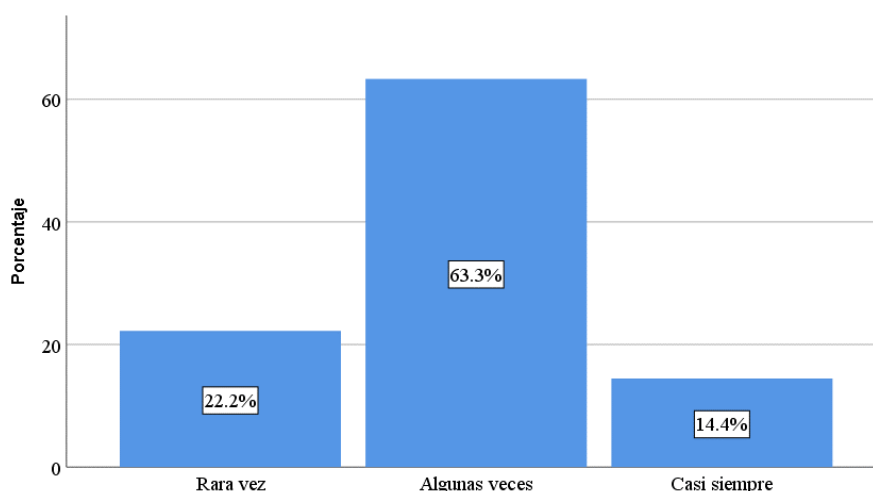
*P06: Asertividad dimensión medio ambiente del algoritmo arboles de decisión*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Rara vez	20	22.2	22.2	22.2
	Algunas veces	57	63.3	63.3	85.6
	Casi siempre	13	14.4	14.4	100.0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 31**

*Po6: Asertividad dimensión medio ambiente del algoritmo arboles de decisión*



*Nota:* Distribución porcentual de la pregunta 05.

Interpretación:

De acuerdo al criterio 06, de asertividad en la dimensión medio ambiente del algoritmo arboles de decisión establece que el 63.3% (57 encuestados) mencionan algunas veces, mientras que el 22.2% (20 encuestados) mencionan rara vez y el 14.4% (13 encuestados) mencionan casi siempre.

**Tabla 28**

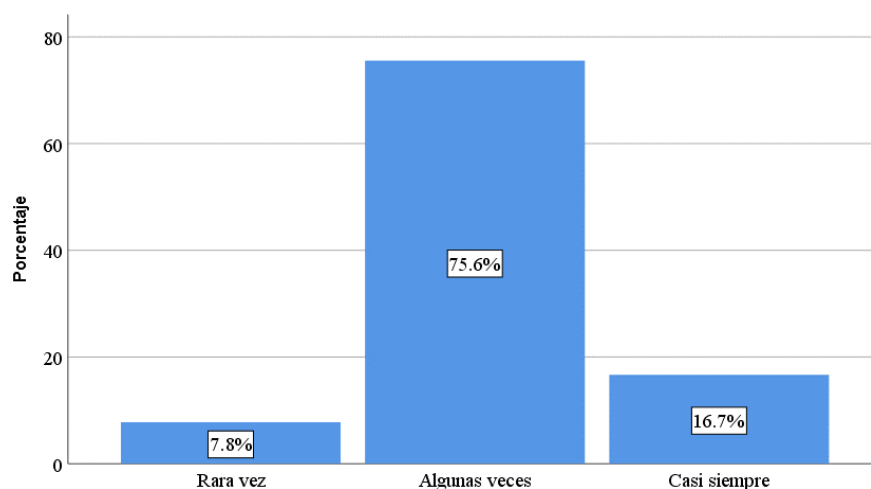
*Po7: Asertividad dimensión instrucción del algoritmo arboles de decisión*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Rara vez	7	7.8	7.8	7.8
	Algunas veces	68	75.6	75.6	83.3
	Casi siempre	15	16.7	16.7	100.0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 32**

**Po7: Asertividad dimensión instrucción del algoritmo arboles de decisión**



*Nota:* Distribución porcentual de la pregunta 07.

Interpretación:

De acuerdo al criterio 07, de asertividad en la dimensión instrucción del algoritmo arboles de decisión establece que el 75.6% (68 encuestados) menciona algunas veces, mientras que el 16.7% (15 encuestados) menciona casi siempre y el 7.8% (7 encuestados) menciona rara.

**Tabla 29**

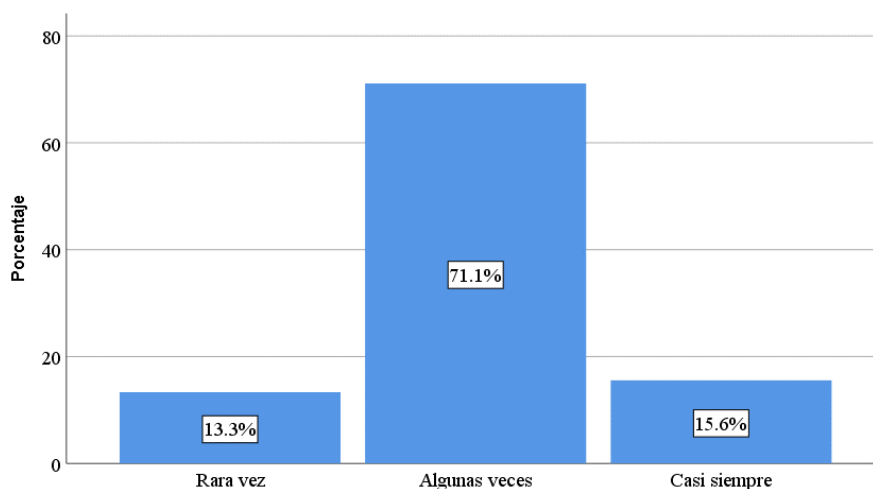
*Po8: Asertividad dimensión logro del algoritmo arboles de decisión*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Rara vez	12	13.3	13.3	13.3
	Algunas veces	64	71.1	71.1	84.4
	Casi siempre	14	15.6	15.6	100.0
	Total	90	100.0	100.0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 33**

*Po8: Asertividad dimensión logro del algoritmo arboles de decisión*



*Nota:* Distribución porcentual de la pregunta o8.

Interpretación:

De acuerdo al criterio o8, de asertividad en la dimensión logro del algoritmo arboles de decisión establece que el 13.3% (12 encuestados) menciona rara vez, mientras que el 71.1% (54 encuestados) menciona algunas veces y el 15.8.8% (14 encuestados) casi siempre.

**Tabla 30**

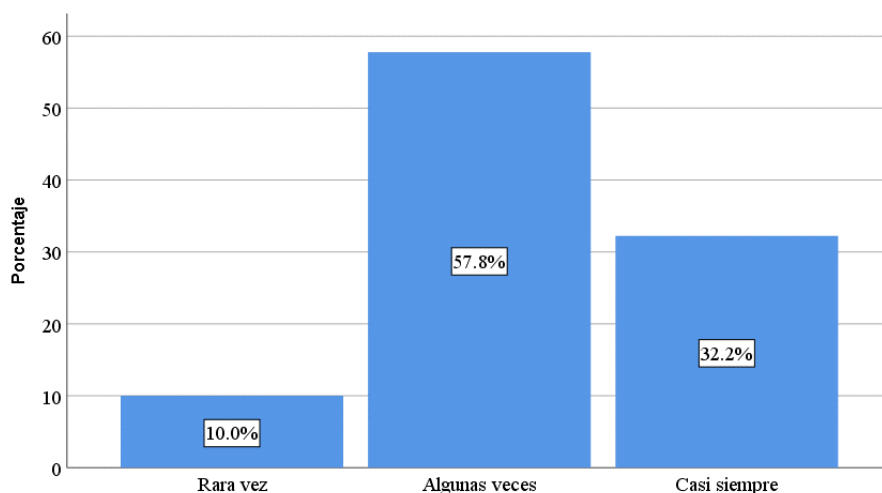
*Po9: Aprendizaje dimensión aptitud del algoritmo Naïve Bayes*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Valido	Rara vez	9	10.0	10.0	10.0
	Algunas veces	52	57.8	57.8	67.8
	Casi siempre	29	32.2	32.2	100.0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 34**

*P09: Aprendizaje dimensión aptitud del algoritmo Naive Bayes*



*Nota:* Distribución porcentual de la pregunta 09.

Interpretación:

De acuerdo al criterio 09, referente al aprendizaje en la dimensión aptitud del algoritmo Naive Bayes establece que el 57.8% (52 encuestados) mencionan algunas veces, mientras que el 32.2% (29 encuestados) mencionan casi siempre y el 10% (9 encuestados) mencionan rara vez.

**Tabla 31**

*P010: Aprendizaje dimensión medio ambiente del algoritmo Naive Bayes*

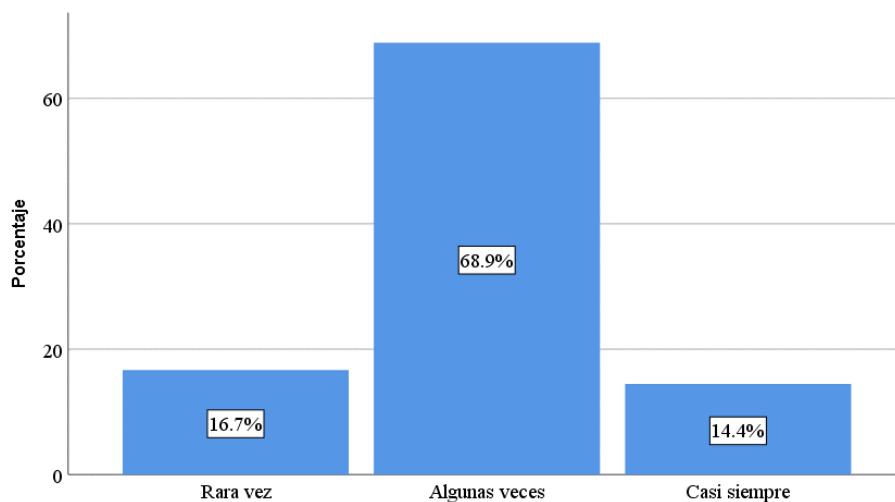
		Frecuencia	Porcentaje	Porcentaje e válido	Porcentaje acumulado
Válido	Rara vez	15	16.7	16.7	16.7
	Algunas veces	62	68.9	68.9	85.6
	Casi siempre	13	14.4	14.4	100.0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.



**Figura 35**

*P010: Aprendizaje dimensión medio ambiente del algoritmo Naive Bayes*



*Nota:* Distribución porcentual de la pregunta 010.

Interpretación:

De acuerdo al criterio 010, referente al aprendizaje en la dimensión medio ambiente del algoritmo Naive Bayes establece que el 68.9% (62 encuestados) mencionan algunas veces, mientras que el 16.7% (15 encuestados) mencionan rara vez y el 14.4% (13 encuestados) mencionan casi siempre.

**Tabla 32**

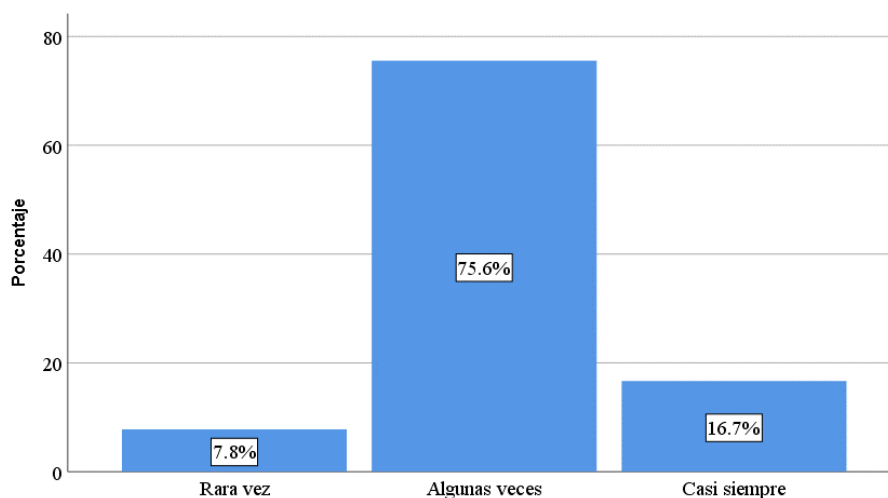
*P011: Aprendizaje dimensión instrucción del algoritmo Naive Bayes*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Rara vez	7	7.8	7.8	7.8
	Algunas veces	68	75.6	75.6	83.3
	Casi siempre	15	16.7	16.7	100.0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 36**

*Po11: Aprendizaje dimensión instrucción del algoritmo Naive Bayes*



*Nota:* Distribución porcentual de la pregunta 011.

Interpretación:

De acuerdo al criterio 011, referente al aprendizaje en la dimensión instrucción del algoritmo Naive Bayes establece que el 75.6% (68 encuestados) mencionan algunas veces, mientras que el 16.7% (15 encuestados) mencionan casi siempre y el 7.8% (7 encuestados) mencionan rara vez.

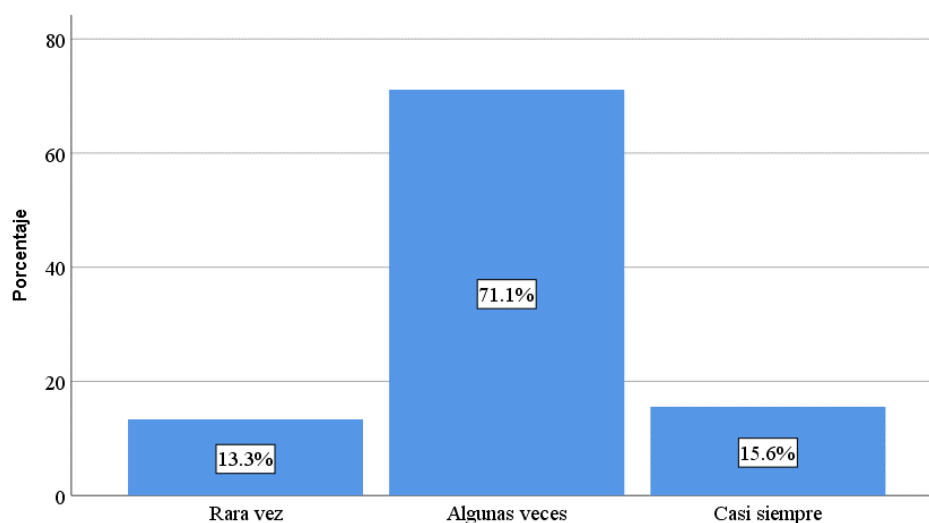
**Tabla 33**

*Po12: Aprendizaje logro del algoritmo Naive Bayes*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Rara vez	12	13.3	13.3	13.3
	Algunas veces	64	71.1	71.1	84.4
	Casi siempre	14	15.6	15.6	100.0
	Total	90	100.0	100.0	

**Figura 37**

*Po12: Aprendizaje logro del algoritmo Naive Bayes*



*Nota:* Distribución porcentual de la pregunta 012.

Interpretación:

De acuerdo al criterio 012, referente al aprendizaje en la dimensión logro del algoritmo Naive Bayes establece que el 13.3% (12 encuestados) mencionan rara vez, mientras que el 71.1% (64 encuestados) mencionan algunas veces y el 15.6% (14 encuestados) mencionan casi siempre.

**Tabla 34**

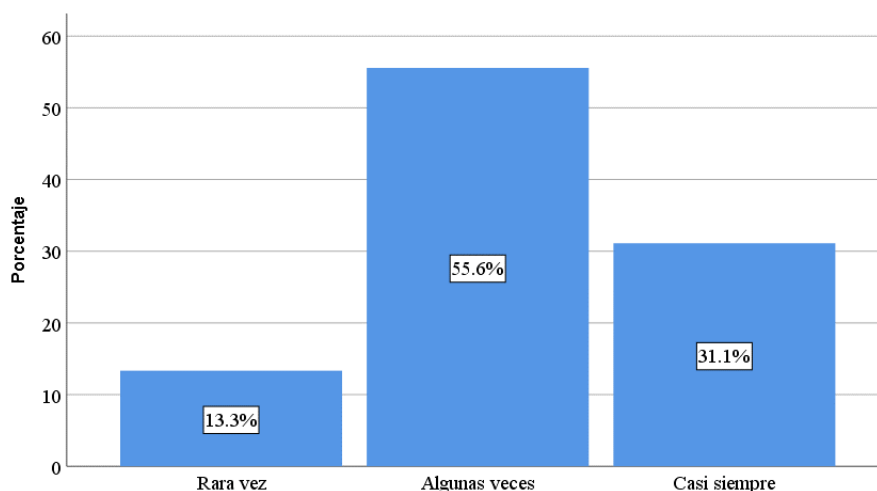
*Po13: Asertividad en la dimensión aptitud del algoritmo Naive Bayes*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Rara vez	12	13.3	13.3	13.3
	Algunas veces	50	55.6	55.6	68.9
	Casi siempre	28	31.1	31.1	100.0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 38**

*P013: Asertividad en la dimensión aptitud del algoritmo Naive Bayes*



*Nota:* Distribución porcentual de la pregunta 013.

Interpretación:

De acuerdo al criterio 013, de asertividad en la dimensión aptitud del algoritmo Naive Bayes establece que el 55.6% (50 encuestados) responde algunas veces, mientras que el 31.1% (28 encuestados) responde casi siempre y el 13.3% (12 encuestados) responde rara vez.

**Tabla 35**

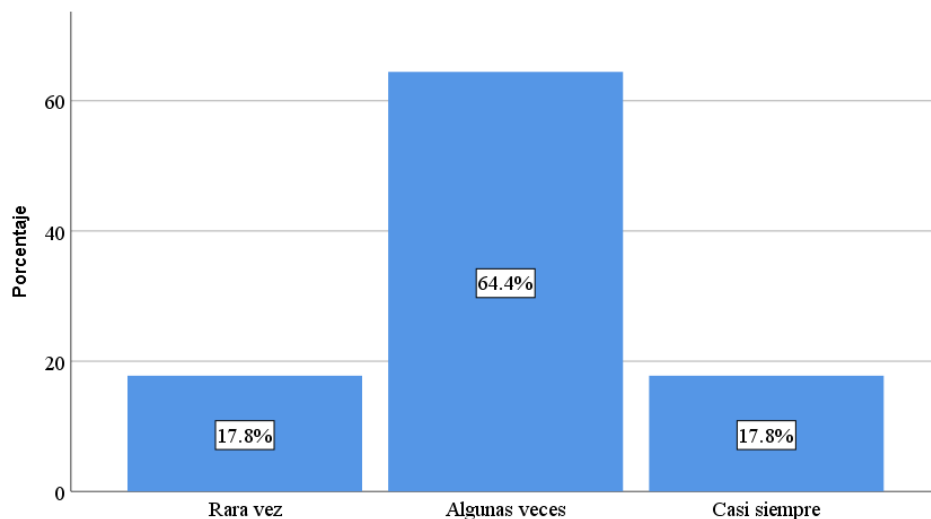
*P014: Asertividad dimensión medio ambiente del algoritmo Naive Bayes*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Rara vez	16	17.8	17.8	17.8
	Algunas veces	58	64.4	64.4	82.2
	Casi siempre	16	17.8	17.8	100.0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 39**

*P014: Asertividad en la dimensión medio ambiente del algoritmo Naive Bayes*



*Nota:* Distribución porcentual de la pregunta 014.

Interpretación:

De acuerdo al criterio 014, de asertividad en la dimensión medio ambiente del algoritmo Naive Bayes establece que el 64.4% (58 encuestados) responde algunas veces, mientras que el 17.8% (16 encuestados) responde rara vez y, casi siempre.

**Tabla 36**

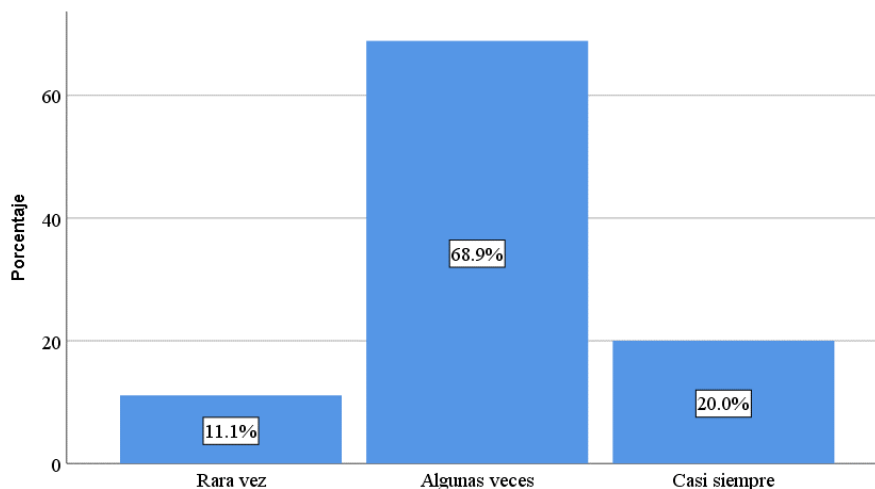
*P015: Asertividad dimensión instrucción del algoritmo Naive Bayes*

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido Rara vez	10	11.1	11.1	11.1
Algunas veces	62	68.9	68.9	80.0
Casi siempre	18	20.0	20.0	100.0
Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 40**

*P015: Asertividad dimensión instrucción del algoritmo Naive Bayes*



*Nota:* Distribución porcentual de la pregunta 015.

Interpretación:

De acuerdo al criterio 015, de asertividad en la dimensión instrucción del algoritmo Naive Bayes establece que el 68.9% (62 encuestados) responde algunas veces, mientras que el 20% (18 encuestados) responde casi siempre y 11.1% (10 encuestados) responde casi siempre.

**Tabla 37**

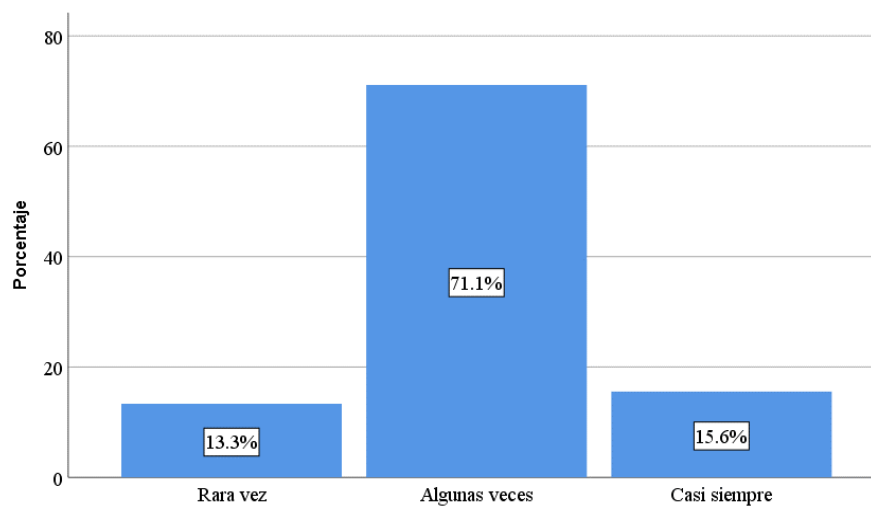
*P016: Asertividad logro del algoritmo Naive Bayes*

			Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Rara vez		12	13.3	13.3	13.3
	Algunas veces		64	71.1	71.1	84.4
	Casi siempre		14	15.6	15.6	100.0
	Total		90	100.0	100.0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 41**

*Po16: Asertividad logro del algoritmo Naive Bayes*



*Nota:* Distribución porcentual de la pregunta 016.

Interpretación:

De acuerdo al criterio 016, referente a asertividad en la dimensión logro del algoritmo Naive Bayes establece que el 13.3% (12 encuestados) mencionan rara vez, mientras que el 71.1% (64 encuestados) mencionan algunas veces y el 15.6% (14 encuestados) mencionan casi siempre.

**Tabla 38**

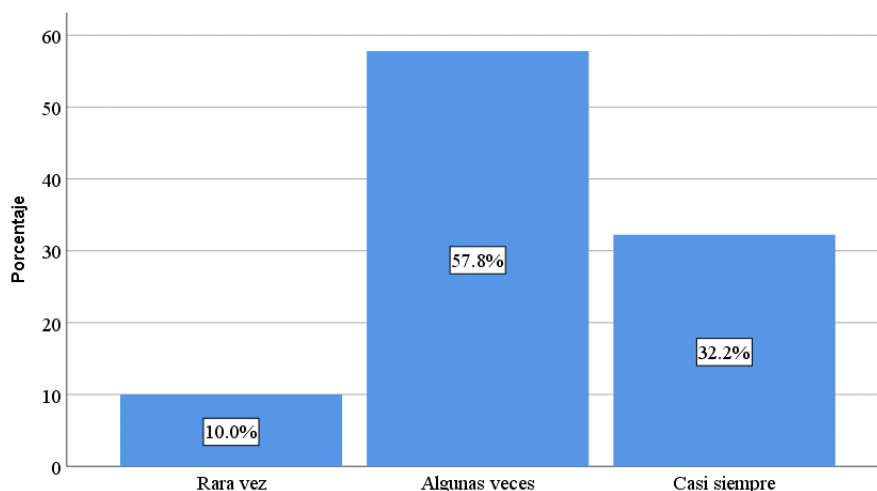
*Po17: Aprendizaje dimensión instrucción del algoritmo Random Forest*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Valido	Rara vez	9	10.0	10.0	10.0
	Algunas veces	52	57.8	57.8	67.8
	Casi siempre	29	32.2	32.2	100.0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 42**

*Po17: Aprendizaje dimensión instrucción del algoritmo Random Forest*



*Nota:* Distribución porcentual de la pregunta 017.

Interpretación:

De acuerdo al criterio 017, referente al aprendizaje en la dimensión instrucción del algoritmo Random Forest establece que el 57.8% (52 encuestados) responden algunas veces, mientras que el 32.2% (29 encuestados) responde casi siempre y el 10% (9 encuestados) mencionan rara vez.

**Tabla 39**

*Po18: Aprendizaje dimensión medio ambiente del algoritmo Random Forest*

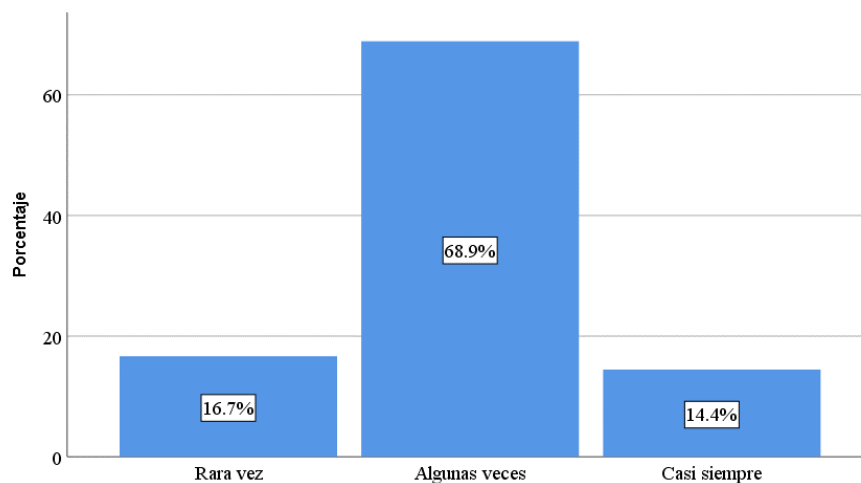
		Frecuencia	Porcentaje	Porcentaj e válido	Porcentaje acumulado
Válido	Rara vez	15	16.7	16.7	16.7
	Algunas veces	62	68.9	68.9	85.6
	Casi siempre	13	14.4	14.4	100.0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.



**Figura 43**

*Po18: Aprendizaje dimensión medio ambiente del algoritmo Random Forest*



*Nota:* Distribución porcentual de la pregunta 018.

Interpretación:

De acuerdo al criterio 018, referente al aprendizaje en la dimensión medio ambiente del algoritmo Random Forest establece que el 68.9% (62 encuestados) responde algunas veces, mientras que el 16.7% (15 encuestados) responde rar vez y el 14.4% (13 encuestados) responde casi siempre.

**Tabla 40**

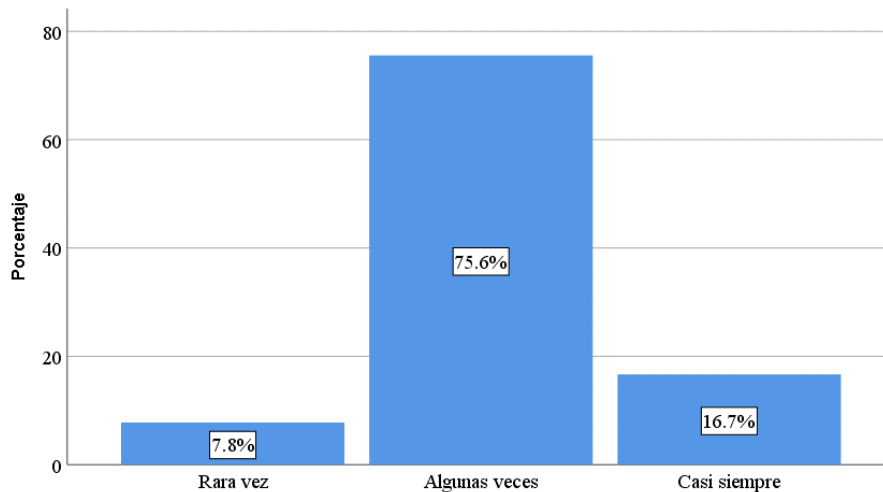
*Po19: Aprendizaje dimensión instrucción del algoritmo Random Forest*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Rara vez	7	7.8	7.8	7.8
	Algunas veces	68	75.6	75.6	83.3
	Casi siempre	15	16.7	16.7	100.0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 44**

*Po19: Aprendizaje dimensión instrucción del algoritmo Random Forest*



*Nota:* Distribución porcentual de la pregunta 019.

Interpretación:

De acuerdo al criterio 019, referente al aprendizaje en la dimensión instrucción del algoritmo Random Forest establece que el 75.6% (68 encuestados) responde algunas veces, mientras que el 16.7% (15 encuestados) responde casi siempre y el 7.8% (7 encuestados) responde rara vez.

**Tabla 41**

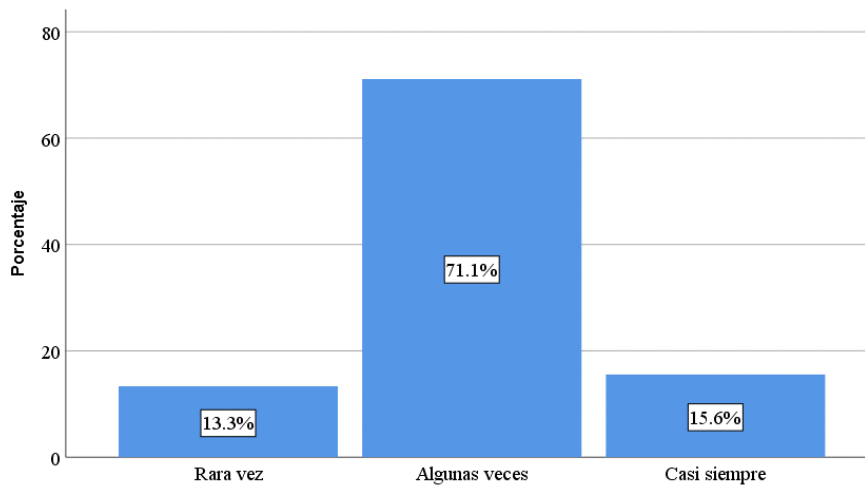
*Po20: Aprendizaje logro del algoritmo Random Forest*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Rara vez	12	13.3	13.3	13.3
	Algunas veces	64	71.1	71.1	84.4
	Casi siempre	14	15.6	15.6	100.0
	Total	90	100.0	100.0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 45**

*Po20: Aprendizaje logro del algoritmo Random Forest*



*Nota:* Distribución porcentual de la pregunta 020.

Interpretación:

De acuerdo al criterio 020, referente a aprendizaje en la dimensión logro del algoritmo Random Forest establece que el 13.3% (12 encuestados) mencionan rara vez, mientras que el 71.1% (64 encuestados) mencionan algunas veces y el 15.6% (14 encuestados) mencionan casi siempre.

**Tabla 42**

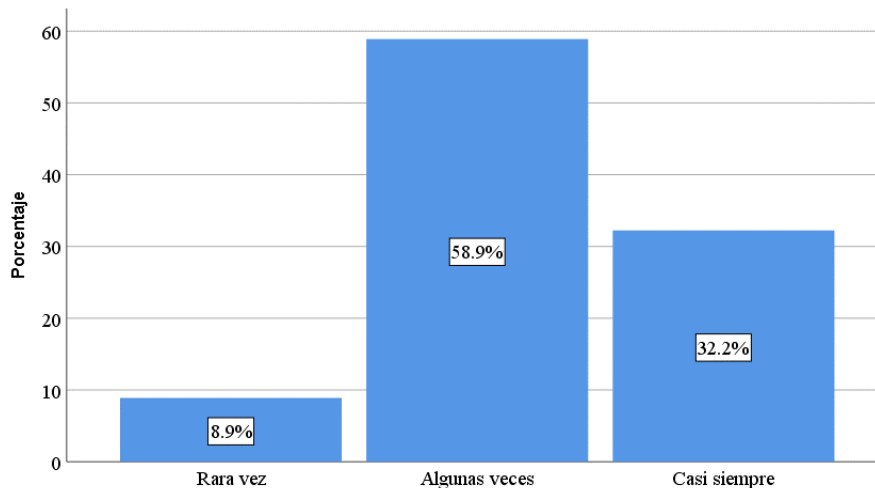
*Po21: Asertividad dimensión aptitud del algoritmo Random Forest*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Rara vez	8	8.9	8.9	8.9
	Algunas veces	53	58.9	58.9	67.8
	Casi siempre	29	32.2	32.2	100.0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 46**

*Po21: Asertividad dimensión aptitud del algoritmo Random Forest*



*Nota:* Distribución porcentual de la pregunta 021.

Interpretación:

De acuerdo al criterio 021, de asertividad en la dimensión aptitud del algoritmo Random Forest establece que el 58.9% (53 encuestados) responde algunas veces, mientras que el 32.2% (29 encuestados) responde casi siempre y el 8.9% (8 encuestados) responde rara vez.

**Tabla 43**

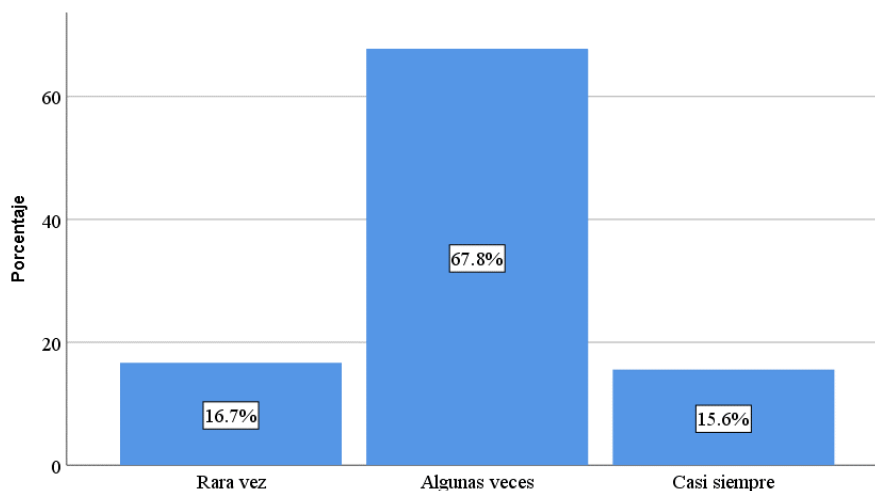
*Po22: Asertividad dimensión medio ambiente del algoritmo Random Forest*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Rara vez	15	16.7	16.7	16.7
	Algunas veces	61	67.8	67.8	84.4
	Casi siempre	14	15.6	15.6	100.0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 47**

*Po22: Asertividad dimensión medio ambiente del algoritmo Random Forest*



*Nota:* Distribución porcentual de la pregunta 022.

Interpretación:

De acuerdo al criterio 022, de asertividad en la dimensión medio ambiente del algoritmo Random Forest establece que el 67.8% (61 encuestados) responde algunas veces, mientras que el 16.7% (15 encuestados) responde rara vez y el 15.6% (14 encuestados) responde casi siempre.

**Tabla 44**

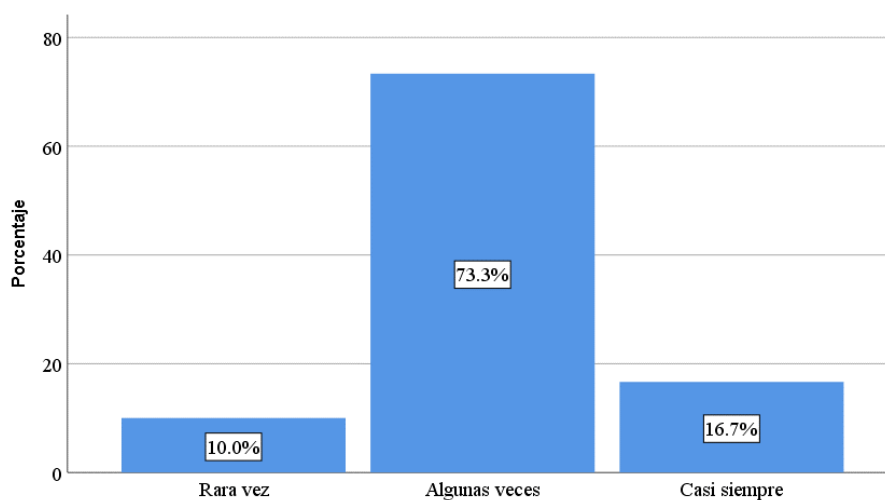
*Po23: Asertividad dimensión instrucción del algoritmo Random Forest*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Rara vez	9	10.0	10.0	10.0
	Algunas veces	66	73.3	73.3	83.3
	Casi siempre	15	16.7	16.7	100.0
	Total	90	100,0	100,0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 48**

*Po23: Asertividad dimensión instrucción del algoritmo Random Forest*



*Nota:* Distribución porcentual de la pregunta 023.

Interpretación:

De acuerdo al criterio 023, de asertividad en la dimensión instrucción del algoritmo Random Forest establece que el 73.3% (66 encuestados) responde algunas veces, mientras que el 16.7% (15 encuestados) responde casi siempre y el 10% (9 encuestados) responde rara vez.

**Tabla 45**

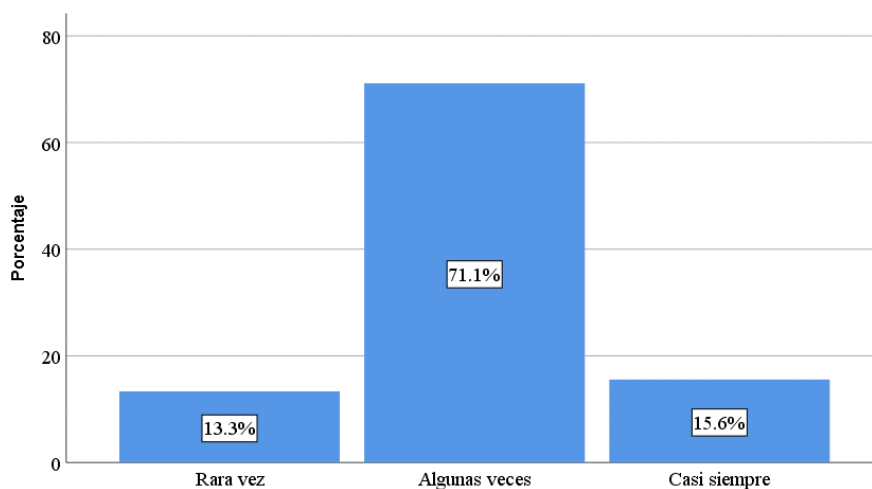
*Po24: Asertividad logro del algoritmo Random Forest*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Rara vez	12	13.3	13.3	13.3
	Algunas veces	64	71.1	71.1	84.4
	Casi siempre	14	15.6	15.6	100.0
	Total	90	100.0	100.0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 49**

*Po24: Asertividad logro del algoritmo Random Forest*



*Nota:* Distribución porcentual de la pregunta 024.

Interpretación:

De acuerdo al criterio 024, referente a la asertividad en la dimensión logro del algoritmo Random Forest establece que el 13.3% (12 encuestados) mencionan rara vez, mientras que el 71.1% (64 encuestados) mencionan algunas veces y el 15.6% (14 encuestados) mencionan casi siempre.

**Tabla 46**

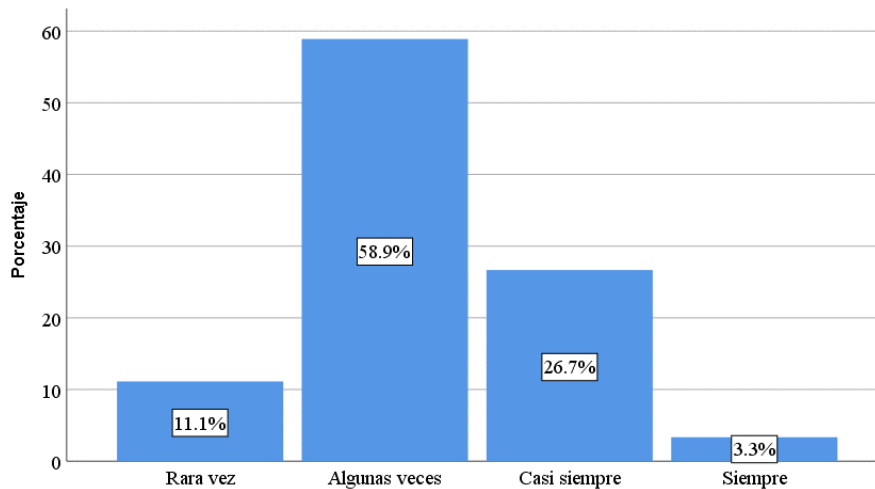
*Po25: Predicción*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Rara vez	10	11.1	11.1	11.1
	Algunas veces	53	58.9	58.9	70.0
	Casi siempre	24	26.7	26.7	96.7
	Siempre	3	3.3	3.3	100.0
	Total	90	100.0	100.0	

*Nota:* Elaboración propia a partir de SPSS versión 26.

**Figura 50**

*P025: Predicción*



*Nota:* Distribución porcentual de la pregunta 025.

Interpretación:

De acuerdo al criterio 025, referente a la predicción de los algoritmos de aprendizaje automático mencionan que el 11.1% (10 encuestados) mencionan rara vez, mientras el 58.9% (53 encuestados) mencionan algunas veces, el 26.7% (24 encuestados) mencionan casi siempre y el 3.3% (3 encuestados) mencionan siempre.

### **Análisis inferencial**

Luego de realizar el análisis del descriptivo de las variables, procederemos a realizar el análisis inferencial que consiste en la prueba de normalidad y prueba de hipótesis planteadas en el presente trabajo de investigación.

### **Prueba de normalidad**

Una herramienta estadística para determinar si un conjunto de datos tiene una distribución normal o gaussiana es la prueba de normalidad (Ghasemi y Zahediasl, 2012), para ello vamos a utilizar la prueba Kolmogorov-Smirnov (K-S), a fin de comprobar si los datos de la muestra proceden de una distribución normal en variables cuantitativas y cuyo tamaño de muestra poblacional sea mayor a 50 (Ghasemi y Zahediasl, 2012; Mishra et al., 2019).



**Tabla 47***Prueba de normalidad*

	Kolmogorov-Smirnov <sup>a</sup>		
	Estadístico	gl	Sig.
Productividad Educativa	.113	90	.006
Algoritmos de Aprendizaje Automático	.143	90	.000

a. Corrección de significación de Lilliefors

*Nota:* Elaboración Propia con Software IBM SPSS versión 24

#### Interpretación:

Se observan en la tabla las pruebas de normalidad para las variables Productividad educativa y Algoritmos de Aprendizaje Automático no siguen una distribución normal, debido a que el p-valor es  $< \alpha$  (0.05) para las dos variables, por lo tanto, utilizaremos pruebas no paramétricas; asimismo, al no pertenecer a una distribución normal se procesó con el Rho de Spearman

#### Prueba no paramétrica con rho de Spemann

Según Martínez Rebollar y Campos Francisco (2015), la prueba de Rho de Spearman es una prueba estadística no paramétrica que se utiliza para medir la asociación directa o inversa entre dos variables cuantitativas (monotónica), la interpretación de este tipo de prueba se basa en el valor de Rho de Spearman, que varía entre -1 y 1.

**Tabla 48***Interpretación del coeficiente de correlación de Spearman*

Valor de Rho	Significado
-1	Correlación negativa grande y perfecta
-0,9 a -0,99	Correlación negativa muy alta
-0,7 a -0,89	Correlación negativa alta
-0,4 a -0,69	Correlación negativa moderada
-0,2 a -0,39	Correlación negativa baja
-0,01 a -0,19	Correlación negativa muy baja

0	Correlación nula
0,01 a 0,19	Correlación positiva muy baja
0,2 a 0,39	Correlación positiva baja
0,4 a 0,69	Correlación positiva moderada
0,7 a 0,89	Correlación positiva alta
0,9 a 0,99	Correlación positiva muy alta
1	Correlación positiva grande y perfecta

*Nota:* (Martínez Rebollar y Campos Francisco, 2015)

### Hipótesis general

H<sub>i</sub>: Existe relación significativa de la productividad educativa con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023.

H<sub>o</sub>: No existe relación significativa de la productividad educativa con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023.

**Tabla 49**

*Productividad Académica y Algoritmos de Aprendizaje Automático.*

			Algoritmos de Aprendizaje			Total
			Automático			
			Insuficiente	Aprobado	Bueno	
Productividad Académica	Insuficiente	Recuento	12	0	0	12
		% del total	13.3%	0.0%	0.0%	13.3%
	Aprobado	Recuento	0	64	0	64
		% del total	0.0%	71.1%	0.0%	71.1%
	Bueno	Recuento	0	0	14	14
		% del total	0.0%	0.0%	15.6%	15.6%
Total	Recuento		12	64	14	90
	% del total		13.3%	71.1%	15.6%	100.0%

*Nota:* Elaboración Propia con Software IBM SPSS versión 26

Interpretación:

El grado de relación existente entre las variables productividad educativa y los algoritmos de aprendizaje automático es 113.3% (12) Insuficiente, 71.1% (64) Aprobado y 15.6% (14) Bueno.

**Tabla 50**

*Prueba no paramétrica. Correlaciones Rho de Spearman entre Productividad Académica y Algoritmos de Aprendizaje Automático.*

			Productividad Académica	Algoritmos de Aprendizaje Automático
Rho de Spearman	Productividad Académica	Coefficiente de correlación	1.000	.940**
		Sig. (bilateral)	.	.000
		N	90	90
	Algoritmos de Aprendizaje Automático	Coefficiente de correlación	.940**	1.000
		Sig. (bilateral)	.000	.
		N	90	90

\*\* . La correlación es significativa en el nivel 0,01 (bilateral).

*Nota:* Elaboración Propia con Software IBM SPSS versión 26

#### Interpretación:

El coeficiente de correlación de Spearman indica una correlación positiva muy alta, y la Rho de Spearman es 0,940. Además, dado que el nivel de significación es  $0,000 < 0,05$ , es posible rechazar la hipótesis nula y aceptar la alternativa, lo que indica que las variables están relacionadas. Esto permite concluir que la productividad académica y los algoritmos de aprendizaje automático están significativamente correlacionados en la EPISI de la UNAM en 2023.

#### Hipótesis específicas

##### Hipótesis específica 1

$H_1$ : Existe relación significativa en la dimensión aptitud con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023.

$H_0$ : No existe relación significativa en la dimensión aptitud con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023.

**Tabla 51***Aptitud y Algoritmos de Aprendizaje Automático.*

			Algoritmos de Aprendizaje Automático			Total
			Insuficiente	Aprobado	Bueno	
Aptitud	Insuficiente	Recuento	7	2	0	9
		% del total	7.8%	2.2%	0.0%	10.0%
	Aprobado	Recuento	5	44	3	52
		% del total	5.6%	48.9%	3.3%	57.8%
	Bueno	Recuento	0	18	11	29
		% del total	0.0%	20.0%	12.2%	32.2%
Total	Recuento		12	64	14	90
	% del total		13.3%	71.1%	15.6%	100.0%

*Nota:* Elaboración Propia con Software IBM SPSS versión 26

Interpretación:

El nivel de relación que existe entre Aptitud y los algoritmos de aprendizaje automático es 7.8% (7) Insuficiente, 48.9% (44) Aprobado y 12.2% (11) Bueno.

**Tabla 52***Prueba no paramétrica. Correlaciones Rho de Spearman entre Aptitud y Algoritmos de Aprendizaje Automático.*

			Aptitud	Algoritmos de Aprendizaje Automático
Rho de Spearman	de Aptitud	Coefficiente de correlación	1.000	.801**
		Sig. (bilateral)	.	.000
		N	90	90
	Algoritmos de Aprendizaje Automático	Coefficiente de correlación	.801**	1.000
		Sig. (bilateral)	.000	.
		N	90	90

\*\*. La correlación es significativa en el nivel 0,01 (bilateral).

*Nota:* Elaboración Propia con Software IBM SPSS versión 26

Interpretación:

Se acepta la hipótesis alternativa y se rechaza la hipótesis nula porque la Rho de Spearman es 0.801, lo que indica una alta correlación positiva por el coeficiente de correlación de Spearman, y porque el nivel de significancia es 0.000, que es menor a 0.05. Esto implica que los factores tienen una relación considerable, por lo que se concluye que en la EPISI de la UNAM en 2023, la aptitud está correlacionada significativamente con los métodos de aprendizaje automático.

#### Hipótesis específica 2

$H_i$ : Existe relación significativa en la dimensión instrucción con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023.

$H_o$ : No existe relación significativa en la dimensión instrucción con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023.

**Tabla 53**

*Instrucción y Algoritmos de Aprendizaje Automático.*

			Algoritmos de Aprendizaje Automático			Total
			Insuficiente	Aprobado	Bueno	
Instrucción	Insuficiente	Recuento	6	1	0	7
		% del total	6.7%	1.1%	0.0%	7.8%
	Aprobado	Recuento	6	54	8	68
		% del total	6.7%	60.0%	8.9%	75.6%
	Bueno	Recuento	0	9	6	15
		% del total	0.0%	10.0%	6.7%	16.7%
Total	Recuento	12	64	14	90	
	% del total	13.3%	71.1%	15.6%	100.0%	

*Nota:* Elaboración Propia con Software IBM SPSS versión 26

#### Interpretación:

El nivel de relación que existe entre Instrucción y los algoritmos de aprendizaje automático es 6.7% (6) Insuficiente, 60% (54) Aprobado y 8.9% (8) Bueno.

**Tabla 54**

*Prueba no paramétrica. Correlaciones Rho de Spearman entre Instrucción y Algoritmos de Aprendizaje Automático.*

			Instrucción	Algoritmos de Aprendizaje Automático
Rho de Spearman	de Instrucción	Coefficiente de correlación	1.000	.723**
		Sig. (bilateral)	.	.000
		N	90	90
	Algoritmos de Aprendizaje Automático	Coefficiente de correlación	.723**	1.000
		Sig. (bilateral)	.000	.
		N	90	90

\*\* . La correlación es significativa en el nivel 0,01 (bilateral).

*Nota:* Elaboración Propia con Software IBM SPSS versión 26

Interpretación:

El valor del coeficiente de correlación de Spearman indica una fuerte asociación positiva, con un Rho de Spearman de 0,723. Además, se acepta la hipótesis alternativa y se rechaza la hipótesis nula porque el nivel de significación es 0.000, que es menor que 0.05. Esto muestra que las variables tienen una asociación significativa, lo que indica que en la EPISI de la UNAM en 2023, la instrucción tiene una relación significativa con los algoritmos de aprendizaje automático.

Hipótesis específica 3

H<sub>i</sub>: Existe relación significativa en la dimensión medio ambiente con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023.

H<sub>0</sub>: No existe relación significativa en la dimensión medio ambiente con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023.

**Tabla 55**

*Medio ambiente y algoritmos de aprendizaje automático.*

Algoritmos de Aprendizaje Automático			Total
Insuficiente	Aprobado	Bueno	

Medio Ambiente	Insuficiente	Recuento	7	8	0	15
		% del total	7.8%	8.9%	0.0%	16.7%
	Aprobado	Recuento	5	53	4	62
		% del total	5.6%	58.9%	4.4%	68.9%
	Bueno	Recuento	0	3	10	13
		% del total	0.0%	3.3%	11.1%	14.4%
Total	Recuento		12	64	14	90
	% del total		13.3%	71.1%	15.6%	100.0
%						

*Nota:* Elaboración Propia con Software IBM SPSS versión 26

Interpretación:

El nivel de relación que existe entre Medio ambiente y los algoritmos de aprendizaje automático es 7.8% (7) Insuficiente, 58.9% (53) Aprobado y 11.1% (10) Bueno.

**Tabla 56**

*Prueba no paramétrica. correlaciones rho de Spearman entre medio ambiente y algoritmos de aprendizaje automático.*

			Medio Ambiente	Algoritmos de Aprendizaje Automático
Rho de Spearman	Medio Ambiente	Coefficiente de correlación	1.000	.626**
		Sig. (bilateral)	.	.000
		N	90	90
	Algoritmos de Aprendizaje Automático	Coefficiente de correlación	.626**	1.000
		Sig. (bilateral)	.000	.
		N	90	90

\*\* . La correlación es significativa en el nivel 0,01 (bilateral).

*Nota:* Elaboración Propia con Software IBM SPSS versión 26

Interpretación:

La correlación moderadamente positiva entre las variables viene indicada por el coeficiente de correlación de Spearman (Rho), que se sitúa en 0,626. Además, dado que el nivel de significancia es de 0.000, es decir, menor a 0.05, se acepta la hipótesis alternativa y se rechaza

la hipótesis nula. Esto indica que, en el EPISI de la UNAM en 2023, habrá una interacción sustancial entre el entorno y los algoritmos de aprendizaje automático.

#### Hipótesis específica 4

$H_i$ : Existe relación significativa en la dimensión aprendizaje con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023.

$H_o$ : No existe relación significativa en la dimensión aprendizaje con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023.

**Tabla 57**

*Aprendizaje y Algoritmos de Aprendizaje Automático.*

			Algoritmos de Aprendizaje Automático			
			Insuficiente	Aprobado	Bueno	Total
Aprendizaje	Insuficiente	Recuento	12	0	0	12
		% del total	13.3%	0.0%	0.0%	13.3%
	Aprobado	Recuento	0	64	0	64
		% del total	0.0%	71.1%	0.0%	71.1%
	Bueno	Recuento	0	0	14	14
		% del total	0.0%	0.0%	15.6%	15.6%
Total	Recuento		12	64	14	90
	% del total		13.3%	71.1%	15.6%	100.0%

*Nota:* Elaboración Propia con Software IBM SPSS versión 26

Interpretación:

El nivel de relación que existe entre aprendizaje y los algoritmos de aprendizaje automático es 13,3% (12) Insuficiente, 71,1% (64) Aprobado y 15,6% (14) Bueno.

**Tabla 58**

*Aprendizaje y Algoritmos de Aprendizaje Automático.*

			Aprendizaje	Algoritmos de Aprendizaje Automático
Rho de	Aprendizaje	Coefficiente de correlación	1.000	.811**
Spearman		Sig. (bilateral)	.	.000



	N	90	90
Algoritmos de Aprendizaje Automático	Coeficiente de correlación	.811**	1.000
	Sig. (bilateral)	.000	.
	N	90	90

\*\* . La correlación es significativa en el nivel 0,01 (bilateral).

*Nota:* Elaboración Propia con Software IBM SPSS versión 26

Interpretación:

Con un coeficiente de correlación de Spearman (Rho) de 0,811, las variables muestran una fuerte asociación positiva entre sí. Además, dado que el nivel de significación es de 0.000, es decir, menor a 0.05, se acepta la hipótesis alternativa y se rechaza la hipótesis nula. Esto indica que, en el EPISI de la UNAM en 2023, los algoritmos de aprendizaje y aprendizaje automático tendrán una relación sustancial.

## Discusión de resultados

En la actualidad, la educación superior se enfrenta a múltiples desafíos, entre ellos, la necesidad de mejorar la productividad educativa y la calidad de la formación de los estudiantes. En este sentido, la implementación de algoritmos de aprendizaje automático (AA) ha sido propuesta como una solución para optimizar el proceso de enseñanza y aprendizaje en la Escuela Profesional de Ingeniería de Sistemas e Informática de la Universidad Nacional de Moquegua. En este estudio, se evaluó la efectividad de la implementación de AA en la productividad educativa de la escuela, es por ello que la presente investigación consiste en determinar el nivel de relación entre la productividad educativa y los algoritmos de aprendizaje automático en la Universidad Nacional de Moquegua, 2023; con la prueba no paramétrica Rho de Spearman se descubrió un valor inferior a 0,05, es decir, un nivel de significación de 0,000. Esto sugiere que las variables tienen una asociación positiva muy fuerte entre sí. Esto sugiere que las variables tienen una asociación positiva muy fuerte entre sí. Por ejemplo, según los datos de 90 estudiantes evaluados en el EPISI de la UNAM, es probable que los elementos de aptitud, entorno e instrucción del constructo de productividad educativa estén significativamente conectados. en donde el 7.8% representado por 7 encuestados estableció insuficiente, el 48.9% representado por 44 encuestados estableció aprobado y el 12.2% representado por 11 encuestados estableció bueno, frente a ello se rechazó la hipótesis nula y aceptando la hipótesis planteada en la investigación referente a la existencia relacional entre la productividad educativa y los algoritmos de aprendizaje automático en EPISI de UNAM, 2023.

Los resultados mencionados confirman la investigación realizada por Parhizkar et al., (2023), denominada “Predicción del desempeño de los estudiantes utilizando algoritmos de clasificación de minería de datos: evaluación de la generalización de modelos desde el aspecto geográfico” Irán, en donde mostró una correlación positiva alta debido a que los algoritmos de aprendizaje automático Random forest y redes neuronales convolucionales (CNN) mostraron mejores rendimientos con un promedio de precisión y una puntuación F de 73,5 y 68,5, respectivamente.

La presente investigación corrobora el estudio "Algoritmos de aprendizaje automático para la predicción del rendimiento académico" de Morales Hernández et al. (2022) realizado en escuelas del estado mexicano de Tlaxcala, en donde se desarrollaron dos clasificadores de AA: el modelo de gradient boosting (GB) y la red neuronal multicapa (también conocida como perceptrón multicapa o MLP, por sus siglas en inglés). El objetivo fue predecir el nivel de productividad académica en las áreas de matemáticas y español. Los resultados demostraron que el clasificador MLP superó al modelo GB en términos de precisión de clasificación general (PG) para la asignatura de español, alcanzando un 70,1% en 2008 y un 61,1% en 2011. Sin embargo, el modelo GB obtuvo mejores resultados en la asignatura de matemáticas, con una PG del 68,8% en 2008 y del 63,5% en 2011. Estos resultados implican que existe una correlación significativa entre el nivel de rendimiento académico en matemáticas y el rendimiento en español.

Asimismo corrobora la investigación realizada por Yağcı (2022), en su modelo de estudio denominado Minería de datos educativos: predicción de la productividad académica de los estudiantes de una universidad estatal turca mediante algoritmos de aprendizaje automático, en donde propone predecir calificaciones de los exámenes finales de los estudiantes de pregrado, en donde referencia las calificaciones de las evaluaciones parciales para dicho estudio, asimismo utiliza algoritmos ML Random Forest, Vecinos más Cercanos, Regresión Logística, Naive Bayes y algoritmos de k-vecino más cercano para dicha investigación, cuyo resultado del modelo propuesto logró una precisión de clasificación del 70% al 75% de relación efectiva positiva.

También lo confirma la investigación realizada por Gismondi (2021) titulado “Modelo predictivo basado en machine learning como soporte para el seguimiento académico del estudiante universitario”, cuya finalidad fue mejorar los resultados en la educación universitaria, proponiendo aplicar la inteligencia artificial, machine learning y Deep learning a través de una red neuronal de seis capas con un 98.97% de precisión en el entrenamiento y

81.73 de precisión en el conjunto de prueba, estableciendo de relación efectiva positiva muy alta.

De la misma manera el estudio realizado por Quiñones y Quiñones (2020)"Rendimiento académico mediante minería de datos" consistió en utilizar técnicas de minería de datos para predecir el rendimiento académico de los alumnos matriculados en la carrera de Ingeniería en Industrias Alimentarias de la Universidad Nacional de Jaén (UNJ). Las oficinas de la universidad y un fichero proporcionaron acceso a la base de datos. CRISP-DM fue la metodología empleada. Los tres algoritmos de categorización del software Weka produjeron predicciones con una fiabilidad superior al 83%.

## **Conclusiones**

1. Conforme al objetivo general en determinar el nivel de relación entre la productividad educativa y los algoritmos de aprendizaje automático, se determinó estadísticamente que existe una correlación positiva muy alta, debido al valor del coeficiente de correlación de Spearman que fue de 0,940 y conforme al baremo de significancia del valor hallado; Por tanto, se destaca que los algoritmos de aprendizaje automático (árboles de decisión, random fores y naive bayes) demostraron un alto grado de precisión en la predicción de la productividad educativa respaldado por el Rho de Spearman.
2. En cuanto al objetivo específico referente al nivel de relación entre la dimensión aptitud y los algoritmos de aprendizaje automático, se determinó estadísticamente que existe una correlación positiva alta debido al valor del coeficiente de correlación de Spearman que fue de 0,801 y el baremo de significancia del valor hallado.
3. En relación al objetivo específico del nivel de relación entre la dimensión instrucción y los algoritmos de aprendizaje automático, se estableció estadísticamente una correlación positiva alta debido al valor del coeficiente de correlación de Spearman que fue de 0,723 y el baremo de significancia del valor hallado.
4. En cuanto al objetivo específico concerniente al nivel de relación entre la dimensión medio ambiente y los algoritmos de aprendizaje automático, se especificó estadísticamente una correlación positiva moderada debido al valor del coeficiente de correlación de Spearman que fue de 0,626 y el baremo de significancia del valor hallado.
5. Respecto al objetivo específico del nivel de relación entre la dimensión aprendizaje y los algoritmos de aprendizaje automático, se analizó estadísticamente en donde se determinó el coeficiente de correlación de Spearman en 0,811 y de acuerdo al baremo de significancia se estableció una correlación positiva alta.

## **Recomendaciones**

1. La recomendación en el ámbito general recae en promover la implementación y uso de algoritmos de aprendizaje automático en el ámbito educativo de EPISI de UNAM, con el objetivo de mejorar la productividad educativa, debido a la alta correlación positiva existente entre dichas variables. Además, recomendar a la vice presidencia académica, decano de ingeniería, jefe de departamento y dirección de EPISI de UNAM brindar el apoyo correspondiente en la asignación de recursos y capacitaciones tanto a estudiantes, docentes y personal involucrado fortaleciendo de esta forma habilidades en el uso y comprensión de los algoritmos de aprendizaje automático (AAA), con la finalidad de lograr mejores niveles de confianza en los AAA, integrar a los sistemas automatizados existentes para una mejor gestión educativa.
2. Con referencia a la dimensión aptitud y los algoritmos de aprendizaje automático, se encontró una correlación positiva alta, por tanto, se recomienda continuar con el fortalecimiento de los datos a fin de mejorar el rendimiento predictivo de los AAA, esto significa incrementar la data historia personalizada, para ello se sugiere a los directivos y autoridades mencionados asignar recursos y capacitaciones a EPISI basado en especialización en inteligencia artificial y Big Data a fin de lograr establecer estrategias y políticas de almacenamiento de datos del modelo inteligente en los servidores de UNAM, para una mejor gestión académica.
3. Respecto a la dimensión instrucción y los algoritmos de aprendizaje automático se obtuvo una correlación alta para lo cual también recomienda ampliar los datos históricos con la finalidad de mejorar las predicciones de los AAA para ello se sugiere a los directivos y autoridades mencionados asignar recursos y capacitaciones a EPISI basado en especialización en inteligencia artificial y Big Data a fin de lograr establecer estrategias y políticas de almacenamiento de datos del modelo inteligente en los servidores de UNAM, para una mejor gestión académica.
4. Respecto a la dimensión medio ambiente y los algoritmos de aprendizaje automático se obtuvo una correlación positiva moderada para lo cual se recomienda mejorar la ampliación de los datos históricos a fin de mejorar las predicciones de los algoritmos de aprendizaje automático para ello se sugiere a los directivos y autoridades mencionados asignar recursos y capacitaciones en Inteligencia artificial y Big Data a fin de establecer estrategias de almacenamiento de datos en los servidores de UNAM, para una mejor gestión académica.

5. Respecto a la dimensión aprendizaje y los algoritmos de aprendizaje automático se obtuvo una correlación positiva alta para lo cual se recomienda mejorar la ampliación de los datos históricos a fin de mejorar las predicciones de los algoritmos de aprendizaje automático para ello se sugiere a los directivos y autoridades mencionados asignar recursos y capacitaciones en Inteligencia artificial y Big Data a fin de establecer estrategias de almacenamiento de datos en los servidores de UNAM, para una mejor gestión académica.

## Referencias

- Arias Gonzáles, J. L. (2021). *Diseño y metodología de la investigación*. Enfoques Consulting E.I.R.L. [https://apps.utel.edu.mx/recursos/files/r161r/w26022w/Arias\\_S2.pdf](https://apps.utel.edu.mx/recursos/files/r161r/w26022w/Arias_S2.pdf)
- Ghasemi, A., y Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, 10(2), 486–489. <https://doi.org/10.5812/ijem.3505>
- Quiñones, L. y Quiñones, Y. L. (2020). Rendimiento académico empleando minería de datos. *Espacios*, 41(44), 277–285. <https://doi.org/10.48082/espacios-a20v41n44p17>
- Martínez Rebollar, A. y Campos Francisco, W. (2015). The Correlation Among Social Interaction Activities Registered Through New Technologies and Elderly's Social Isolation Level. *Revista Mexicana de Ingeniería Biomédica*, 36(3), 177–188. <https://doi.org/10.17488/RMIB.36.3.4>
- Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., & Keshri, A. (2019). Application of Student's t-test, analysis of variance, and covariance in research. *Annals of Cardiac Anaesthesia*, 22(4), 407–411. [https://doi.org/10.4103/aca.ACA\\_94\\_19](https://doi.org/10.4103/aca.ACA_94_19)
- Morales Hernández, M. Á. y González Camacho, J. M. y Robles Vásquez, H. y Del Valle Paniagua, D. H. y Durán Moreno, J. R. (2022). Algoritmos de aprendizaje automático para la predicción del logro académico. *RIDE Revista Iberoamericana Para La Investigación y El Desarrollo Educativo*, 12(24). <https://doi.org/10.23913/ride.v12i24.1180>
- Parhizkar, A. y Tejeddin, G. y Khatibi, T. (2023). Student performance prediction using datamining classification algorithms: Evaluating generalizability of models from geographical aspect. *Education and Information Technologies*
- Yağcı, M. (2022). Educational data mining: prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments*, 9(1), 11. <https://doi.org/10.1186/s40561-022-00192-z>

## Referencias generales.

Abid, N. y Ali, R. y Akhter, M. (2021). Exploring gender-based difference towards academic enablers scales among secondary school students of Pakistan. *Psychology in the Schools*, 58(7), 1380–1398. <https://doi.org/10.1002/pits.22538>

Ahmed Khan, Z. y Adnan, J. y Adnan Raza, S. (2023). *Cognitive Learning Theory and Development: Higher Education Case Study*. <https://doi.org/10.5772/intechopen.110629>

Al-Ababneh M. (2020). Linking Ontology, Epistemology and Research Methodology. *Science & Philosophy*.

Al-Ababneh, M. M. (2020). Linking Ontology, Epistemology and Research Methodology. *Science & Philosophy*, 75–91.

Al-Ani, O. y Das, S. (2022). Reinforcement Learning: Theory and Applications in hems. *Mathematics & computer science, Artificial Intelligence & Robotics* .

Alban, J. y Calero J. L. (2017). El rendimiento académico: aproximación necesaria a un problema pedagógico actual. *Revista Conrado*.

Alkin, M. C. y King, J. A. (2016). The Historical Development of Evaluation Use. *American Journal of Evaluation*, 37(4), 568–579. <https://doi.org/10.1177/1098214016665164>

Andrade, M. S. y Freitas, J. C. de. (2023). Analysis of Performance Metrics on the Conjunction of Intrusions in IEEE 802.11 Networks with Machine Learning at Hospital N.S.C. In *connecting expertise multidisciplinary development for the future*. Seven Editora. <https://doi.org/10.56238/Connexpemultidisdevoipfut-116>

Anyoha, R. (2017). *La historia de la inteligencia artificial - Ciencia en las noticias*. <https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/>

Auffarth, B. (2021). *Machine Learning for Time-Series with Python*. [https://books.google.com.pe/books?id=a7tLEAAQBAJ&pg=PA99&lpg=PA99&dq=author+of+nearest+neighbors+algorithm++Evelyn+Fix+y+Joseph+Hodges&source=bl&ots=D19ck8Ao7W&sig=ACfU3UovGNfCvt\\_kI7tpjETbQgZOoGuyAA&hl=es&sa=X&ved=2ahUKEwiTt474wcj4AhVVAdQKHUSABNUQ6AF6BAgoEAM#v=onepage&q=author%20of%20nearest%20neighbors%20algorithm%20%20Evelyn%20Fix%20y%20Joseph%20Hodges&f=false](https://books.google.com.pe/books?id=a7tLEAAQBAJ&pg=PA99&lpg=PA99&dq=author+of+nearest+neighbors+algorithm++Evelyn+Fix+y+Joseph+Hodges&source=bl&ots=D19ck8Ao7W&sig=ACfU3UovGNfCvt_kI7tpjETbQgZOoGuyAA&hl=es&sa=X&ved=2ahUKEwiTt474wcj4AhVVAdQKHUSABNUQ6AF6BAgoEAM#v=onepage&q=author%20of%20nearest%20neighbors%20algorithm%20%20Evelyn%20Fix%20y%20Joseph%20Hodges&f=false)

Avila, Á. (2022). *La motivación en los estudiantes: 3 Claves para aumentarla*. <https://www.enriccorberainstitute.com/blog/motivacion-en-los-estudiantes/>

Bagnato Juan Ignacio. (2020). *Aprende Machine Learning en Español Teoría + Práctica Python*. 1–184.

Bernal Torres, C. A. (2016). *Metodología de la investigación: Administración, economía, humanidades y ciencias sociales* (4<sup>a</sup> ed.). Pearson Educación. <https://drive.google.com/file/d/1-3wqx7vGGCn6O4FxMPkzKw15E4tByYXX/view?pli=1>

Barus, S. P. (2021). Implementation of Naïve Bayes Classifier-based Machine Learning to Predict and Classify New Students at Matana University. *Journal of Physics: Conference Series*, 1842(1), 012008. <https://doi.org/10.1088/1742-6596/1842/1/012008>

BBVA. (2019). “Machine learning”: ¿qué es y cómo funciona?

Biblioteca Nacional de Chile. (2021). *La Filosofía Positivista*. <http://www.memoriachilena.gob.cl/602/w3-article-93966.html>

Brew, E. A. y Nketiah, B. y Koranteng, R. (2021). A Literature Review of Academic Performance, an Insight into Factors and their Influences on Academic Outcomes of Students at Senior High Schools. *OALib*, 08(06), 1–14. <https://doi.org/10.4236/oalib.1107423>

Brownlee, J. (2018). *Machine Learning Algorithms From Scratch* (v.1.8, Ed.).

Carroll, J. B. (1989). The Carroll Model: A 25-Year Retrospective and Prospective View. *Educational Researcher*, 18(1), 26. <https://doi.org/10.2307/1176007>

Carrasco Díaz, S. (2006). *Metodología de la investigación científica: pautas metodológicas para diseñar y elaborar el proyecto de investigación*. San Marcos.

<https://drive.google.com/file/d/1GTWMTyAZDmzEohJbUKSxsR-QJWsYugBV/view>

Caselli Gismondi, H. E. (2021). *Modelo predictivo basado en machine learning como soporte para el seguimiento académico del estudiante universitario* [Universidad Nacional del Santa]. <http://repositorio.uns.edu.pe/bitstream/handle/UNS/3804/52337.pdf?sequence=5&isAllowed=y>

Caselli Gismondi, H. E. y Urrelo Huiman, L. V. (2021). Características para un modelo de predicción de la deserción académica universitaria. Caso Universidad Nacional de Santa. *Llamkasun*, 2(4), 02–22. <https://doi.org/10.47797/llamkasun.v2i4.61>

Castañón, E. R. (2023). *¿Qué es un ambiente de aprendizaje según autores?* <https://www.centrobanamex.com.mx/que-es-un-ambiente-de-aprendizaje-segun-autores/>

Castrillón, O. D. y Sarache, W. y Ruiz-Herrera, S. (2020). Predicción del rendimiento académico por medio de técnicas de inteligencia artificial. *Formación Universitaria*, 13(1), 93–102. <https://doi.org/10.4067/S0718-50062020000100093>

Cecilia, O. N. y Bernedette, C.-U. y Emmanuel, A. E. y Hope, A. N. (2020). Enhancing students academic performance in Chemistry by using kitchen resources in Ikom, Calabar. *Educational Research and Reviews*, 15(1), 19–26. <https://doi.org/10.5897/ERR2019.3810>

Charbuty, B. y Abdulazeez, A. (2021). Classification Based on Decision Tree Algorithm for Machine Learning. *Journal of Applied Science and Technology Trends*, 2(01), 20–28. <https://doi.org/10.38094/jastt20165>

Contreras, L. E. y Fuentes, H. J. y Rodríguez, J. I. (2020). Predicción del rendimiento académico como indicador de éxito/fracaso de los estudiantes de ingeniería, mediante aprendizaje automático. *Formación Universitaria*, 13(5), 233–246. <https://doi.org/10.4067/S0718-50062020000500233>

Correia, M. I. T. D. (2023). Ethics in research. *Clinical Nutrition Open Science*, 47, 121–130. <https://doi.org/10.1016/j.nutos.2022.12.010>

de Oliveira Chagas, E. T. (2019). Deep Learning and its applications today. *Revista Científica Multidisciplinar Núcleo Do Conhecimento*, 04(05), 05–26. <https://doi.org/10.32749/nucleodoconhecimento.com.br/business-administration/deep-learning-2>

Díaz-Landa, B. y Meleán-Romero, R. y Marín-Rodríguez, W. (2021). Rendimiento académico de estudiantes en Educación Superior: predicciones de factores influyentes a partir de árboles de decisión. *Telos Revista de Estudios Interdisciplinarios En Ciencias Sociales*, 23(3), 616–639. <https://doi.org/10.36390/telos233.08>

Dong, Y. y Hou, J. y Zhang, N. y Zhang, M. (2020). Research on How Human Intelligence, Consciousness, and Cognitive Computing Affect the Development of Artificial Intelligence. *Complexity*, 2020, 1–10. <https://doi.org/10.1155/2020/1680845>

Edgard, T. W. y Manz, D. O. (2017). *Research Methods for Cyber Security* (1ra ed.). Elsevier Inc.

Escobar-Pérez, J. y Cuervo-Martínez, Á. (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. *Avances En Medición*, 27–36. [https://www.humanas.unal.edu.co/lab\\_psicometria/application/files/9416/0463/3548/Vol\\_6.\\_Articulo3\\_Juicio\\_de\\_expertos\\_27-36.pdf](https://www.humanas.unal.edu.co/lab_psicometria/application/files/9416/0463/3548/Vol_6._Articulo3_Juicio_de_expertos_27-36.pdf)

Esteso, A. y Peidro, D. y Mula, J. y Díaz-Madroñero, M. (2022). Reinforcement learning applied to production planning and control. *International Journal of Production Research*, 1–18. <https://doi.org/10.1080/00207543.2022.2104180>

Fernandes, E. y Holanda, M. y Victorino, M. y Borges, V. y Carvalho, R. y Erven, G. van. (2019). Educational data mining: Predictive analysis of academic performance of public school students in the capital of Brazil. *Journal of Business Research*, 94, 335–343. <https://doi.org/10.1016/j.jbusres.2018.02.012>



Fernández-Mellizo, M. y Constante-Amores, A. (2020). Determinantes del rendimiento académico de los estudiantes de nuevo acceso a la Universidad Complutense de Madrid. *Revista de Educación*, 213–240.

Figuerola-Abarzúa, C. y Meza-Vásquez, S. y Estrada-Lagos, R. (2022). Desarrollo de capacidad argumentativa en estudiantes universitarios, mediante uso del debate como estrategia didáctica. *South Florida Journal of Development*, 3(6), 6328–6346. <https://doi.org/10.46932/sfjdv3n6-001>

Fu, M. y Zhang, C. y Hu, C. y Wu, T. y Dong, J. y Zhu, L. (2023). Achieving Verifiable Decision Tree Prediction on Hybrid Blockchains. *Entropy*, 25(7), 1058. <https://doi.org/10.3390/e25071058>

Géron, A. (2019). *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow Concepts, Tools, and Techniques to Build Intelligent System* (Ni. Tache, Ed.; 2nd ed.).

Ghasemi, A. y Zahediasl, S. (2012). Normality Tests for Statistical Analysis: A Guide for Non-Statisticians. *International Journal of Endocrinology and Metabolism*, 10(2), 486–489. <https://doi.org/10.5812/ijem.3505>

Gogus, A. (2012). Bloom's Taxonomy of Learning Objectives. In *Encyclopedia of the Sciences of Learning* (pp. 469–473). Springer US. [https://doi.org/10.1007/978-1-4419-1428-6\\_141](https://doi.org/10.1007/978-1-4419-1428-6_141)

Gopani, A. (2022). *La historia de los algoritmos de aprendizaje automático*. <https://analyticsindiamag.com/the-history-of-machine-learning-algorithms/>

Grasso Imig, P. (2020). *Rendimiento académico: un recorrido conceptual que aproxima a una definición unificada para el ámbito superior*. [https://fh.mdp.edu.ar/revistas/index.php/r\\_educ/article/view/4165](https://fh.mdp.edu.ar/revistas/index.php/r_educ/article/view/4165)

Guill, K. y Ömeroğulları, M. y Köller, O. (2021). Intensity and content of private tutoring lessons during German secondary schooling: effects on students' grades and test achievement. *European Journal of Psychology of Education*. <https://doi.org/10.1007/s10212-021-00581-x>

Hajjej, F. y Alohalı, M. A. y Badr, M. y Rahman, M. A. (2022). A Comparison of Decision Tree Algorithms in the Assessment of Biomedical Data. *BioMed Research International*, 2022, 1–9. <https://doi.org/10.1155/2022/9449497>

Hathcoat, J. D. y Meixner, C. y Nicholas, M. C. (2019). Ontology and Epistemology. In *Handbook of Research Methods in Health Social Sciences* (pp. 99–116). Springer Singapore. [https://doi.org/10.1007/978-981-10-5251-4\\_56](https://doi.org/10.1007/978-981-10-5251-4_56)

Hawkins, C. y Bailey, L. E. (2020). A New Data Landscape: IR's Role in Academic Analytics. *New Directions for Institutional Research*, 2020(185–186), 87–103. <https://doi.org/10.1002/ir.20331>

Hernández Sampieri, R. y Fernández Collado, C. y Baptista Lucio, M. del P. (2014). *Metodología de la investigación*. Mc Graw Hill.

Hernandez-Manxilla J.M. (2011). El racionalismo cartesiano y su particular conquista de la subjetividad en el mundo moderno. *FRENIA*.

Hernández-Sampieri R. (2018). *Metodología de la investigación* (1ra Ed.).

Herrera Acosta, C. E. y Sánchez Pinedo, L. D. (2019). METODOLOGÍA DE APRENDIZAJE Y SISTEMA DE EVALUACIÓN PARA ALCANZAR RESULTADOS EN EL PROCESO EDUCATIVO . *Evaluación de La Calidad Educativa*. <https://www.igobernanza.org/index.php/IGOB/Contacto>

Hiller, J. (2016). *Epistemological foundations of objectivist and interpretivist research* (pp. 99–127).

IBM Cloud Education. (2020). *Machine Learning This introduction to machine learning provides an overview of its history, important definitions, applications, and concerns within businesses today*. <https://www.ibm.com/cloud/learn/machine-learning>

Ilić, M. y Srdjević, Z. y Srdjević, B. (2022). Water quality prediction based on Naïve Bayes algorithm. *Water Science and Technology*, 85(4), 1027–1039. <https://doi.org/10.2166/wst.2022.006>

Janiesch, C. y Zschech, P. y Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3), 685–695. <https://doi.org/10.1007/s12525-021-00475-2>

Jiang, T. y Gradus, J. L. y Rosellini, A. J. (2020). Supervised Machine Learning: A Brief Primer. *Behavior Therapy*, 51(5), 675–687. <https://doi.org/10.1016/j.beth.2020.05.002>

Jin, S. y Fang, G. y Cheung, K. C. y Sit, P. S. (2022). Factors associated with academic resilience in disadvantaged students: An analysis based on the PISA 2015 B-S-J-G (China) sample. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.846466>

Jovel, J. y Greiner, R. (2021). An Introduction to Machine Learning Approaches for Biomedical Research. *Frontiers in Medicine*, 8. <https://doi.org/10.3389/fmed.2021.771607>

Kaun, C. y Jhanjhi, N. Z. y Goh, W. W. y Sukumaran, S. (2021). Implementation of Decision Tree Algorithm to Classify Knowledge Quality in a Knowledge Intensive System. *MATEC Web of Conferences*, 335, 04002. <https://doi.org/10.1051/mateconf/202133504002>

Kaur Arora, S. (2022). *Top Steps To Learn Naive Bayes Algorithm*. <https://hackr.io/blog/top-steps-to-learn-naive-bayes-algorithm>

KOÇOĞLU, F. Ö. (2022). Research on the success of unsupervised learning algorithms in indoor location prediction. *International Advanced Researches and Engineering Journal*, 6(2 (under construction)), 148–153. <https://doi.org/10.35860/iarej.1096573>

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. In *Theory Into Practice* (Vol. 41, Issue 4). [https://doi.org/10.1207/s15430421tip4104\\_2](https://doi.org/10.1207/s15430421tip4104_2)

Lampropoulos, A. S. y Tsihrintzis, G. A. (2015). *Machine Learning Paradigms* (Vol. 92). Springer International Publishing. <https://doi.org/10.1007/978-3-319-19135-5>

Laplane, P. A. (2001). *Dictionary of computer science, engineering, and technology*. [www.crepress.com](http://www.crepress.com)

Lee, J. y Warner, E. y Shaikhouni, S. y Bitzer, M. y Kretzler, M. y Gipson, D. y Pennathur, S. y Belovich, K. y Bhat, Z. y Gadegbeku, C. y Massengill, S. y Perumal, K. y Saha, J. y Yang, Y. y Luo, J. y Zhang, X. y Mariani, L. y Hodgins, J. B. y Rao, A. (2022). Unsupervised machine learning for identifying important visual features through bag-of-words using histopathology data from chronic kidney disease. *Scientific Reports*, 12(1), 4832. <https://doi.org/10.1038/s41598-022-08974-8>

Mahesh, B. (2018). Machine Learning Algorithms - A Review. *International Journal of Science and Research*, 381–386.

Maina, J. J. y Zakari, A. T. y Alkali, I. A. y Salisu, R. A. (2021). Academic success predictors for architecture students at Kano University of Science and Technology, Wudil, Kano State, Nigeria. *Bayero Journal of Pure and Applied Sciences*, 13(2), 125–133. <https://doi.org/10.4314/bajopas.v13i2.17>

Mandal, L. y Jana, N. D. (2019). A Comparative Study of Naive Bayes and k-NN Algorithm for Multi-class Drug Molecule Classification. *2019 IEEE 16th India Council International Conference (INDICON)*, 1–4. <https://doi.org/10.1109/INDICON47234.2019.9029095>

Martínez Rebollar, A. y Campos Francisco, W. (2015). The Correlation Among Social Interaction Activities Registered Through New Technologies and Elderly's Social Isolation Level. *Revista Mexicana de Ingeniería Biomédica*, 36(3), 177–188. <https://doi.org/10.17488/RMIB.36.3.4>

Mazana, M. Y. y Montero, C. S. y Casimir, R. O. (2019). Investigating Students' Attitude towards Learning Mathematics. *International Electronic Journal of Mathematics Education*, 14(1). <https://doi.org/10.29333/iejme/3997>

McDaniel, M. y Storey, V. C. (2020). Evaluating Domain Ontologies. *ACM Computing Surveys*, 52(4), 1–44. <https://doi.org/10.1145/3329124>

McGinnis, D. (2020). *¿Qué es la Cuarta Revolución Industrial? | Fuerza de ventas.* <https://www.salesforce.com/blog/what-is-the-fourth-industrial-revolution-4ir/>

Medina N. y Ferreira J. y Marzol R. (2018). Factores personales que inciden en el bajo rendimiento académico de los estudiantes de geometría. *Revista de Estudios Interdisciplinarios En Ciencias Sociales.*

Menacho Chiok, C. H. (2017). Predicción del rendimiento académico aplicando técnicas de minería de datos. *Anales Científicos*, 78(1), 26. <https://doi.org/10.21704/ac.v78i1.811>

Mienye, I. D. y Sun, Y. y Wang, Z. (2019). Prediction performance of improved decision tree-based algorithms: a review. *Procedia Manufacturing*, 35, 698–703. <https://doi.org/10.1016/j.promfg.2019.06.011>

Mishra, P. y Pandey, C. y Singh, U. y Gupta, A. y Sahu, C. y Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Annals of Cardiac Anaesthesia*, 22(1), 67. [https://doi.org/10.4103/aca.ACA\\_157\\_18](https://doi.org/10.4103/aca.ACA_157_18)

Misra, S. y Li, H. (2020). Noninvasive fracture characterization based on the classification of sonic wave travel times. In *Machine Learning for Subsurface Characterization* (pp. 243–287). Elsevier. <https://doi.org/10.1016/B978-0-12-817736-5.00009-0>

Mohammed, M. y Khan, M. B. y Mohammed Bashier, E. B. (2017). *Machine Learning Algorithms and Applications.*

Morales Hernández, M. Á. y González Camacho, J. M. y Robles Vásquez, H. y Del Valle Paniagua, D. H. y Durán Moreno, J. R. (2022). Algoritmos de aprendizaje automático para la predicción del logro académico. *RIDE Revista Iberoamericana Para La Investigación y El Desarrollo Educativo*, 12(24). <https://doi.org/10.23913/ride.v12i24.1180>

Münch, M. y Raab, C. y Biehl, M. y Schleif, F.-M. (2020). Data-Driven Supervised Learning for Life Science Data. *Frontiers in Applied Mathematics and Statistics*, 6. <https://doi.org/10.3389/fams.2020.553000>

Nedeva, V. y Pehlivanova, T. (2021). Students' Performance Analyses Using Machine Learning Algorithms in WEKA. *IOP Conference Series: Materials Science and Engineering*, 1031(1), 012061. <https://doi.org/10.1088/1757-899X/1031/1/012061>

Ong, A. K. S. y Chuenyindee, T. y Prasetyo, Y. T. y Nadlifatin, R. y Persada, S. F. y Gumasing, Ma. J. J. y German, J. D. y Robas, K. P. E. y Young, M. N. y Sittiwatethanasiri, T. (2022). Utilization of Random Forest and Deep Learning Neural Network for Predicting Factors Affecting Perceived Usability of a COVID-19 Contact Tracing Mobile Application in Thailand “ThaiChana.” *International Journal of Environmental Research and Public Health*, 19(10), 6111. <https://doi.org/10.3390/ijerph19106111>

Orlikowsld W. y Baroudi J. (1991). Studying Information Technology in Organimtions: Research Approaches and Assumptions. *Information Systems Research*.

Osisanwo, F. Y. y Akinsola, J. E. T. y Awodele, O. y Hinmikaiye, J. O. y Olakanmi, O. y Akinjobi, J. (2017). Supervised Machine Learning Algorithms: Classification and Comparison. *International Journal of Computer Trends and Technology*, 48(3), 128–138. <https://doi.org/10.14445/22312803/IJCTT-V48P126>

Oxford University. (2016). *A Dictionary of Computer Science*.

Padilla-Cuevas, J. y Reyes-Ortiz, J. A. y Bravo, M. (2021). Ontology-Based Context Event Representation, Reasoning, and Enhancing in Academic Environments. *Future Internet*, 13(6), 151. <https://doi.org/10.3390/fi13060151>

Palanichamy, K. (2019). Integrative Omic Analysis of Neuroblastoma. In *Computational Epigenetics and Diseases* (pp. 311–326). Elsevier. <https://doi.org/10.1016/B978-0-12-814513-5.00019-2>

Parhizkar, A. y Tejeddin, G. y Khatibi, T. (2023). Student performance prediction using datamining classification algorithms: Evaluating generalizability of models from geographical aspect. *Education and Information Technologies*, 28(11), 14167–14185. <https://doi.org/10.1007/s10639-022-11560-0>

Park, Y. S. y Konge, L. y Artino, A. R. (2020). The Positivism Paradigm of Research. *Academic Medicine*, 95(5), 690–694. <https://doi.org/10.1097/ACM.0000000000003093>

Parth Shukla. (2023). *Naive Bayes Algorithms: A Complete Guide for Beginners*. <https://www.analyticsvidhya.com/blog/2023/01/naive-bayes-algorithms-a-complete-guide-for-beginners/>

Pastor, M. (1997). *Deep Blue venció a Kasparov en un campo donde el hombre no tenía rival* | | *ComputerWorld*. Computer World. <https://www.computerworld.es/archive/deep-blue-vencio-a-kasparov-en-un-campo-donde-el-hombre-no-tenia-rival>

Peconcillo Jr, L. B. y D. Peteros, E. y O. Mamites, I. y T. Sanchez, D. y L. Tenerife, J. J. y L. Suson, R. (2020). Structuring Determinants to Level Up Students Performance. *International Journal of Education and Practice*, 8(4), 638–651. <https://doi.org/10.18488/journal.61.2020.84.638.651>

Pellegrino, E. y Jacques, C. y Beaufils, N. y Nanni, I. y Carliz, A. y Metellus, P. y Ouafik, L. (2021). Machine learning random forest for predicting oncosomatic variant NGS analysis. *Scientific Reports*, 11(1), 21820. <https://doi.org/10.1038/s41598-021-01253-y>

Pirneskoski, J. y Tamminen, J. y Kallonen, A. y Nurmi, J. y Kuisma, M. y Olkkola, K. T. y Hoppu, S. (2020). Random forest machine learning method outperforms prehospital National

Early Warning Score for predicting one-day mortality: A retrospective study. *Resuscitation Plus*, 4, 100046. <https://doi.org/10.1016/j.resplu.2020.100046>

Prabhakaran, S. (2018). *How Naive Bayes Algorithm Works?* Machine Learning. <https://www.machinelearningplus.com/predictive-modeling/how-naive-bayes-algorithm-works-with-example-and-full-code/>

Pugliese, R. y Regondi, S. y Marini, R. (2021). Machine learning-based approach: global trends, research directions, and regulatory standpoints. *Data Science and Management*, 4, 19–29. <https://doi.org/10.1016/j.dsm.2021.12.002>

Quiñones, L. y Quiñones, Y. L. (2020). Rendimiento académico empleando minería de datos. *Espacios*, 41(44), 277–285. <https://doi.org/10.48082/espacios-a20v41n44p17>

Ray, S. (2019). A Quick Review of Machine Learning Algorithms. *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, 35–39. <https://doi.org/10.1109/COMITCon.2019.8862451>

Raynor, W. J. (1999). *The International Dictionary of Artificial Intelligence*.

Rigdon, J. C. (2016). *Dictionary of Computer and Internet Terms* (Microsoft Corporation, Ed.; 1st Edition). <http://www.wordsrus.info>

Roman, V. (2019). *Algoritmos Naive Bayes: Fundamentos e Implementación*. Ciencia de Datos. <https://medium.com/datos-y-ciencia/algoritmos-naive-bayes-fudamentos-e-implementaci%C3%B3n-4bcb24b307f>

Rosales Sánchez, E. M. y Rodríguez Ortega, P. G. y Romero Ariza, M. (2020). Conocimiento, demanda cognitiva y contextos en la evaluación de la alfabetización científica en PISA. *Revista Eureka Sobre Enseñanza y Divulgación de Las Ciencias*, 17(2), 1–22. [https://doi.org/10.25267/Rev\\_Eureka\\_ensen\\_divulg\\_cienc.2020.v17.i2.2302](https://doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2020.v17.i2.2302)

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 19–27.

Russell, S. y Norvig, P. (2008). *Inteligencia Artificial Un Enfoque Moderno* (D. F. Aragón, Ed.; 2nd ed.).

Salatino, A. A. y Thanapalasingam, T. y Mannocci, A. y Birukou, A. y Osborne, F. y Motta, E. (2020). The Computer Science Ontology: A Comprehensive Automatically-Generated Taxonomy of Research Areas. *Data Intelligence*, 2(3), 379–416. [https://doi.org/10.1162/dint\\_a\\_00055](https://doi.org/10.1162/dint_a_00055)

Sarker, I. H. (2021a). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, 2(3), 160. <https://doi.org/10.1007/s42979-021-00592-x>

Sarker, I. H. (2021b). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, 2(3), 160. <https://doi.org/10.1007/s42979-021-00592-x>

Sauce, B. y Liebherr, M. y Judd, N. y Klingberg, T. (2022). The impact of digital media on children's intelligence while controlling for genetic differences in cognition and socioeconomic background. *Scientific Reports*, 12(1), 7720. <https://doi.org/10.1038/s41598-022-11341-2>

Saunders M. y Lewis, P. y Thornhill, A. (2019). *Research Methods for Business Students* (Pearson Education, Ed.; Eighth).

Saunders, M. N. K. y Lewis, P. y Thornhill, A. (2029). *Research Methods for Business Students* (8va ed.).

Schonlau, M. y Zou, R. Y. (2020). The random forest algorithm for statistical learning. *The Stata Journal: Promoting Communications on Statistics and Stata*, 20(1), 3–29. <https://doi.org/10.1177/1536867X20909688>

Scriven, Mi. (1967). *The methodology of evaluation*.

Seel, N. M. (2012a). Bloom's Model of School Learning. In *Encyclopedia of the Sciences of Learning* (pp. 466–469). Springer US. [https://doi.org/10.1007/978-1-4419-1428-6\\_979](https://doi.org/10.1007/978-1-4419-1428-6_979)

Seel, N. M. (2012b). Carroll's Model of School Learning. In *Encyclopedia of the Sciences of Learning* (pp. 501–503). Springer US. [https://doi.org/10.1007/978-1-4419-1428-6\\_980](https://doi.org/10.1007/978-1-4419-1428-6_980)

Sepúlveda, A. y Minte, A. y Díaz-Levicoy, D. (2020). Caracterización de preguntas en libros de texto de Ciencias Naturales en Educación Primaria chilena. *Educação e Pesquisa*, 46. <https://doi.org/10.1590/s1678-4634202046224118>

Serumena D. y Utan F. y Poernomo M. (2021). The Effectiveness of Social Media as An Online Learning Pattern in Improving the 3 Domains of Student Intellectual Ability During the Pandemic (Covid-19). *Advances in Engineering Research*.

Sharma, N. y Sharma, R. y Jindal, N. (2021). Machine Learning and Deep Learning Applications-A Vision. *Global Transitions Proceedings*, 2(1), 24–28. <https://doi.org/10.1016/j.gltp.2021.01.004>

Simplilearn. (2022). *The Complete Guide to Machine Learning Steps*. <https://www.simplilearn.com/tutorials/machine-learning-tutorial/machine-learning-steps>

Singh, V. y Chen, S.-S. y Singhania, M. y Nanavati, B. y kar, A. kumar y Gupta, A. (2022). How are reinforcement learning and deep learning algorithms used for big data based decision making in financial industries—A review and research agenda. *International Journal of Information Management Data Insights*, 2(2), 100094. <https://doi.org/10.1016/j.jjime.2022.100094>

Sivasubramaniam, S. y Dlabolová, D. H. y Kralikova, V. y Khan, Z. R. (2021). Assisting you to advance with ethics in research: an introduction to ethical governance and application procedures. *International Journal for Educational Integrity*, 17(1), 14. <https://doi.org/10.1007/s40979-021-00078-6>

Stern, M. y Arinze, C. y Perez, L. y Palmer, S. E. y Murugan, A. (2020). Supervised learning through physical changes in a mechanical system. *Proceedings of the National Academy of Sciences*, 117(26), 14843–14850. <https://doi.org/10.1073/pnas.2000807117>

Tai, M.-T. (2020). The impact of artificial intelligence on human society and bioethics. *Tzu Chi Medical Journal*, 32(4), 339. [https://doi.org/10.4103/tcmj.tcmj\\_71\\_20](https://doi.org/10.4103/tcmj.tcmj_71_20)

Tamayo y Tamayo, M. (2004). *El proceso de la investigación científica*. Limusa. [https://www.academia.edu/29308889/Tamayo\\_Mario\\_El\\_Proceso\\_De\\_La\\_Investigacion\\_Cientifica\\_pdf](https://www.academia.edu/29308889/Tamayo_Mario_El_Proceso_De_La_Investigacion_Cientifica_pdf)

Tebani, A. y Bekri, S. (2020). High-throughput omics in the precision medicine ecosystem. In *Precision Medicine for Investigators, Practitioners and Providers* (pp. 19–31). Elsevier. <https://doi.org/10.1016/B978-0-12-819178-1.00003-4>

Torres-Malca, J. R. y Vera-Ponce, V. J. y Zuzunaga-Montoya, F. E. y Talavera, J. E. y De La Cruz-Vargas, J. A. (2022). Content validity by expert judgment of an instrument to measure knowledge, attitudes and practices about salt consumption in the peruvian population. *Revista de La Facultad de Medicina Humana*, 22(2), 273–279. <https://doi.org/10.25176/RFMH.v22i2.4768>

Triwiyanto T. y Suyanto y Prasojo L.D. y Wardana Y. (2020). Factors Affecting Educational Productivity at Private Elementary Schools in Indonesia. *Advances in Social Science, Education and Humanities Research*, 487, 318–323.

Universidad de Lehigh. (2013). *Historia de las Pruebas Estandarizadas*. <https://ed.lehigh.edu/news-events/news/history-standardized-testing>

Vadlamudi, P. S. y Gunasekaran, M. y Nagalakshmi, T. J. (2023). An Analysis of the Effectiveness of the Naive Bayes Algorithm and the Support Vector Machine for Detecting Fake News on Social Media. *2023 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE)*, 726–731. <https://doi.org/10.1109/IITCEE57236.2023.10090978>

Vigo, R. y Zeigler, D. E. y Wimsatt, J. (2022). Uncharted Aspects of Human Intelligence in Knowledge-Based “Intelligent” Systems. *Philosophies*, 7(3), 46. <https://doi.org/10.3390/philosophies7030046>

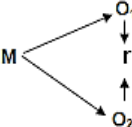


- Villalba, F. (2018). Naive Bayes- clasificación bayesiano ingenuo. In *Aprendizaje supervisado en R*. <https://fervilber.github.io/Aprendizaje-supervisado-en-R/ingenuo.html>
- Vostroknutov, A. y Polonio, L. y Coricelli, G. (2018). The Role of Intelligence in Social Learning. *Scientific Reports*, 8(1), 6896. <https://doi.org/10.1038/s41598-018-25289-9>
- Walberg, H. J. (1984). *Improving the Productivity of America's Schools*.
- Wang, P. (2019). On Defining Artificial Intelligence. *Journal of Artificial General Intelligence*, 10(2), 1–37. <https://doi.org/10.2478/jagi-2019-0002>
- Xiang, X. y Foo, S. y Zang, H. (2021). Recent Advances in Deep Reinforcement Learning Applications for Solving Partially Observable Markov Decision Processes (POMDP) Problems Part 2—Applications in Transportation, Industries, Communications and Networking and More Topics. *Machine Learning and Knowledge Extraction*, 3(4), 863–878. <https://doi.org/10.3390/make3040043>
- Yağcı, M. (2022). Educational data mining: prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments*, 9(1), 11. <https://doi.org/10.1186/s40561-022-00192-z>
- Yeturu, K. (2020). *Machine learning algorithms, applications, and practices in data science* (pp. 81–206). <https://doi.org/10.1016/bs.host.2020.01.002>
- Yildiz, M. y Börekci, C. (2020). Predicting Academic Achievement with Machine Learning Algorithms. *Journal of Educational Technology and Online Learning*. <https://doi.org/10.31681/jetol.773206>
- Zambrano Aranda, G. (2022). Análisis de la satisfacción con respecto a una maestría de una universidad de Lima desde la perspectiva de sus egresados. *Revista Educación y Sociedad*, 3(5), 9–22. <https://doi.org/10.53940/reys.v3i5.89>
- Zhao, W. (2022). Inspired, but not mimicking: a conversation between artificial intelligence and human intelligence. *National Science Review*, 9(6). <https://doi.org/10.1093/nsr/nwac068>
- Zheng, Y. (2019). Identification of microRNAs From Small RNA Sequencing Profiles. In *Computational Non-coding RNA Biology* (pp. 35–82). Elsevier. <https://doi.org/10.1016/B978-0-12-814365-0.00012-9>
- Žukauskas, P. y Vveinhardt, J. y Andriukaitienė, R. (2018). Philosophy and Paradigm of Scientific Research. In *Management Culture and Corporate Social Responsibility*. InTech. <https://doi.org/10.5772/intechopen.70628>

## **ANEXOS**

## ANEXO 1: MATRIZ DE CONSISTENCIA

Productividad educativa y algoritmos de aprendizaje automático en escuela de ingeniería de sistemas e informática, universidad nacional de Moquegua, 2023

Problema	Objetivos	Hipótesis	Variables	Metodología
<p><b>Problema general:</b> ¿Cuál es el nivel de relación que existe entre la productividad educativa y los algoritmos de aprendizaje en estudiantes EPISI de UNAM, 2023?</p> <p><b>Problemas específicos:</b> ¿Cuál es el nivel de relación que existe entre la dimensión aptitud y los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023?</p> <p>¿Cuál es el nivel de relación que existe entre la dimensión instrucción y los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023?</p>	<p><b>Objetivo general</b> Determinar el nivel de relación que existe entre la productividad educativa y los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023</p> <p><b>Objetivos específicos:</b> Determinar el nivel de relación que existe entre la dimensión aptitud y los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023</p> <p>Establecer el nivel de relación que existe entre la dimensión instrucción y los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023</p>	<p><b>Hipótesis General</b> Existe relación significativa de la productividad educativa con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023</p> <p><b>Hipótesis específicas</b> Existe relación significativa en la dimensión aptitud con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023</p> <p>Existe relación significativa en la dimensión instrucción con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023</p>	<p><b>Variable 1</b> Rendimiento académico</p> <p><b>Dimensión</b> Aptitud</p> <p><b>Indicador:</b> – capacidad (1,2) – motivación (3,4) – desarrollo (5,6)</p> <p><b>Dimensión</b> Instrucción</p> <p><b>Indicador</b> – Cantidad (14,15, 16) – Calidad (17, 18)</p> <p><b>Dimensión</b> Medio ambiente</p> <p><b>Indicador</b> – hogar (7, 8, 9) – entorno comunitario (10, 11, 12) – medios de comunicación (13)</p> <p><b>Dimensión</b> Aprendizaje</p> <p><b>Indicador</b> – Logro (19)</p> <p><b>Variable 2:</b> Algoritmos de aprendizaje automático</p> <p><b>Dimensión</b> Algoritmo de Árboles de Decisión</p>	<p><b>Enfoque:</b> Cuantitativo</p> <p><b>Tipo de Investigación:</b> Básica</p> <p><b>Nivel de Investigación:</b> Descriptivo Correlacional</p> <p><b>Método de Investigación:</b> Hipotético deductivo</p> <p><b>Diseño de la investigación:</b> No experimental</p>  <p>Donde: M = Muestra O<sub>1</sub> = Observación de la V.1. O<sub>2</sub> = Observación de la V.2. r = Correlación entre dichas variables.</p> <p><b>Población:</b> 116 estudiantes</p> <p><b>Variable 1:</b></p> <p><b>Muestra:</b> 90 estudiantes</p> <p><b>Técnica:</b> Encuesta</p> <p><b>Instrumento:</b> Cuestionario</p> <p><b>Variable 2:</b></p> <p><b>Muestra:</b> 90 estudiantes</p> <p><b>Técnica:</b> Observación</p> <p><b>Instrumento:</b> Lista de cotejo</p>

Problema	Objetivos	Hipótesis	Variables	Metodología
UNAM, 2023?  ¿Cuál es el nivel de relación que existe entre la dimensión medio ambiente y los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023?  ¿Cuál es el nivel de relación que existe entre la dimensión aprendizaje y los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023?	Especificar el nivel de relación que existe entre la dimensión medio ambiente y los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023  Analizar el nivel de relación que existe entre la dimensión aprendizaje y los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023	Existe relación significativa en la dimensión medio ambiente con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023  Existe relación significativa en la dimensión de aprendizaje con los algoritmos de aprendizaje automático en estudiantes EPISI de UNAM, 2023	Indicadores – Aprendizaje (01,02,03) – Asertividad (04, 05, 06)  Dimensión Algoritmo Naive bayes  Indicadores – Aprendizaje (07,08,09) – Asertividad (010, 011, 012)  Dimensión Algoritmo de Random Forest  Indicadores – Aprendizaje (013,014, 015) – Asertividad (016, 017, 018)  Dimensión Aprendizaje automático  Indicador – Predicción (019)	

## ANEXO 2: INSTRUMENTOS DE RECOLECCION DE DATOS

Variable 1: Productividad Educativa

### ENCUESTA

**Indicación:** Por favor conteste el presente cuestionario según criterio solicitado.

Con el motivo de desarrollar el trabajo de investigación “PRODUCTIVIDAD EDUCATIVA Y ALGORITMOS DE APRENDIZAJE AUTOMATICO EN ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA, UNAM, 2023” se solicita y agradece su colaboración e indicarle, que la presente encuesta es totalmente confidencial, cuyos resultados se darán a conocer únicamente en forma tabulada e impersonal es fundamental que sus respuestas sean fundamentadas en la verdad,

**Objetivo.** Recoger información necesaria para promover el desarrollo de la tecnología inteligente en la Universidad y así poder identificar el desarrollo de la tecnología y su aplicación de conocimientos.

**Información específica:** Lea los ítems que se citan a continuación y marque con (X) el casillero de su preferencia.

1	2	3	4	5
Nunca	Rara Vez	Algunas Veces	Casi Siempre	Siempre

**Escuela Profesional:** Ingeniería de Sistemas e Informática

**Sexo:** \_\_\_\_\_

**Edad:** \_\_\_\_\_

Preguntas	1	2	3	4	5
Dimensión Aptitud					
1. Utilizo los recursos académicos disponibles para mis conocimientos					
2. Planifico mis estudios académicos en la universidad					
3. Los docentes felicitan mis esfuerzos académicos					
4. Comprendo las sesiones de aprendizaje					

Preguntas	1	2	3	4	5
5. Planifico mis trabajos académicos					
6. Utilizo la biblioteca para ampliar mis conocimientos					
Dimensión Medio Ambiente					
7. Mis padres apoyan en mis conocimientos académicos					
8. Utilizo equipos y herramientas académicas de casa para estudiar					
9. Recibo apoyo financiero familiar para mis estudios					
10. Realizo desenvolvimiento social con mis vecinos					
11. Participo en actividades deportivas en mi barrio					
12. Participo en actividades comunitarias con los conocimientos adquiridos					
13. Los medios de comunicación apoyan el desarrollo de mis conocimientos					
Dimensión Instrucción					
14. Planifico mis actividades académicas constantemente					
15. Planifico el uso de TICs para mis estudios e investigaciones					
16. Planifico realizar proyectos de emprendimiento académico					
17. Existe socialización académica con los docentes					
18. Existe satisfacción académica en mi experiencia universitaria					
DIMENSIÓN APRENDIZAJE					
19. El nivel de logro de aprendizaje es satisfactorio					

Variable 2: Algoritmos de Aprendizaje Automático

### LISTA DE COTEJO

Escuela Profesional: \_\_\_\_\_

Nombre algoritmo aprendizaje automático: \_\_\_\_\_

Periodo de evaluación: \_\_\_\_\_ Fecha: \_\_\_\_\_

Instrucción:

Con el motivo de desarrollar el trabajo de investigación “PRODUCTIVIDAD EDUCATIVA Y ALGORITMOS DE APRENDIZAJE AUTOMATICO EN ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA, UNAM, 2023”, verificar que se encuentren los componentes señalados abajo y marque con un aspa (X) el registro de cumplimiento correspondiente, en caso de ser necesario hay un espacio de observaciones ´ para realizar la retroalimentación.

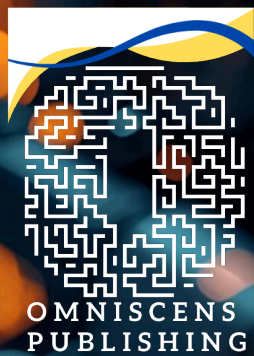
**Información específica:** Lea los ítems que se citan, se marcan con (X) el casillero.

1	2	3	4	5
Nunca	Rara Vez	Algunas Veces	Casi Siempre	Siempre

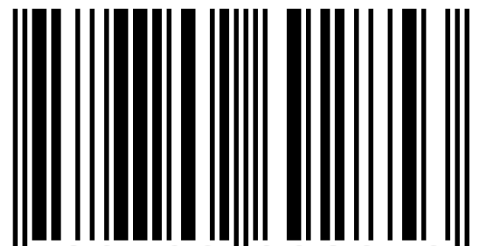
Ítems	1	2	3	4	5
Dimensión Algoritmos Arboles de decisión					
01. Aprendizaje Aptitud					
02. Aprendizaje Medio ambiente					
03. Aprendizaje instrucción					
04. Aprendizaje Logro					
05. Asertividad Aptitud					
06. Asertividad Medio ambiente					
07. Asertividad instrucción					
08. Asertividad Logro					
Dimensión Algoritmo Naive Bayes					
09. Aprendizaje Aptitud					
10. Aprendizaje Medio ambiente					
11. Aprendizaje instrucción					
12. Aprendizaje Logro					
13. Asertividad Aptitud					

14. Asertividad Medio ambiente					
15. Asertividad instrucción					
16. Asertividad Logro					
Dimensión Algoritmo Random Forest					
17. Aprendizaje Aptitud					
18. Aprendizaje Medio ambiente					
19. Aprendizaje instrucción					
20. Aprendizaje Logro					
21. Asertividad Aptitud					
22. Asertividad Medio ambiente					
23. Asertividad instrucción					
24. Asertividad Logro					
Dimensión Aprendizaje Automático					
25. Predicción					





ISBN: 978-970-96928-5-3



9 789709 692853