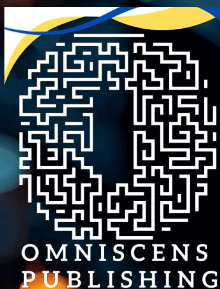


CAPÍTULO 4

Hipótesis, variables y metodología de la investigación



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



CAPÍTULO 4

Hipótesis, variables y metodología de la investigación

Hypothesis, variables and research methodology

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

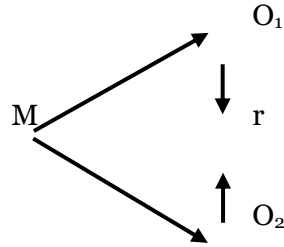
Métodos y diseño de investigación.

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
 O₁ = Observación de la V1
 O₂ = 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

SI² = La varianza del instrumento

ST² = 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
- 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
- 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
- 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>