



**UNIVERSIDAD CÉSAR VALLEJO**

**FACULTAD DE INGENIERÍA Y ARQUITECTURA**  
**ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS**

Aplicación Web utilizando árboles de decisión para el rendimiento académico del curso de Comunicación - I.E. 20955-24 Jorge Guevara Mellado

**TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE:**

Ingeniera de Sistemas

**AUTORAS:**

Ramon Vitteri, Mariana (orcid.org/0000-0001-7816-0301)  
Velasquez Espinoza, Becky Kazumy (orcid.org/0000-0003-2481-6293)

**ASESOR:**

Dr. Sanchez Atuncar, Giancarlo (orcid.org/0000-0001-9842-7317)

**LÍNEA DE INVESTIGACIÓN:**

Sistemas de Información y Comunicaciones

**LÍNEA DE RESPONSABILIDAD SOCIAL UNIVERSITARIA:**

Desarrollo económico, empleo y emprendimiento

LIMA – PERÚ

2024



**UNIVERSIDAD CÉSAR VALLEJO**

**FACULTAD DE INGENIERÍA Y ARQUITECTURA**

**ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS**

**Declaratoria de Autenticidad del Asesor**

Yo, SANCHEZ ATUNCAR GIANCARLO, docente de la FACULTAD DE INGENIERÍA Y ARQUITECTURA de la escuela profesional de INGENIERÍA DE SISTEMAS de la UNIVERSIDAD CÉSAR VALLEJO SAC - LIMA ATE, asesor de Tesis titulada: "Aplicación Web utilizando árboles de decisión para el rendimiento académico del curso de Comunicación - I.E. 20955-24 Jorge Guevara Mellado", cuyos autores son RAMON VITTERI MARIANA, VELASQUEZ ESPINOZA BECKY KAZUMY, constato que la investigación tiene un índice de similitud de 17%, verificable en el reporte de originalidad del programa Turnitin, el cual ha sido realizado sin filtros, ni exclusiones.

He revisado dicho reporte y concluyo que cada una de las coincidencias detectadas no constituyen plagio. A mi leal saber y entender la Tesis cumple con todas las normas para el uso de citas y referencias establecidas por la Universidad César Vallejo.

En tal sentido, asumo la responsabilidad que corresponda ante cualquier falsedad, ocultamiento u omisión tanto de los documentos como de información aportada, por lo cual me someto a lo dispuesto en las normas académicas vigentes de la Universidad César Vallejo.

LIMA, 04 de Julio del 2024

<b>Apellidos y Nombres del Asesor:</b>	<b>Firma</b>
SANCHEZ ATUNCAR GIANCARLO <b>DNI:</b> 41488834 <b>ORCID:</b> 0000-0001-9842-7317	Firmado electrónicamente por: GSANCHEZAT el 04- 07-2024 13:16:23

Código documento Trilce: TRI - 0793192





**Declaratoria de Originalidad de los Autores**

Nosotros, RAMON VITTERI MARIANA, VELASQUEZ ESPINOZA BECKY KAZUMY estudiantes de la FACULTAD DE INGENIERÍA Y ARQUITECTURA de la escuela profesional de INGENIERÍA DE SISTEMAS de la UNIVERSIDAD CÉSAR VALLEJO SAC - LIMA ATE, declaramos bajo juramento que todos los datos e información que acompañan la Tesis titulada: "Aplicación Web utilizando árboles de decisión para el rendimiento académico del curso de Comunicación - I.E. 20955-24 Jorge Guevara Mellado", es de nuestra autoría, por lo tanto, declaramos que la Tesis:

1. No ha sido plagiada ni total, ni parcialmente.
2. Hemos mencionado todas las fuentes empleadas, identificando correctamente toda cita textual o de paráfrasis proveniente de otras fuentes.
3. No ha sido publicada, ni presentada anteriormente para la obtención de otro grado académico o título profesional.
4. Los datos presentados en los resultados no han sido falseados, ni duplicados, ni copiados.

En tal sentido asumimos la responsabilidad que corresponda ante cualquier falsedad, ocultamiento u omisión tanto de los documentos como de la información aportada, por lo cual nos sometemos a lo dispuesto en las normas académicas vigentes de la Universidad César Vallejo.

Nombres y Apellidos	Firma
RAMON VITTERI MARIANA <b>DNI:</b> 70570204 <b>ORCID:</b> 0000-0001-7816-0301	Firmado electrónicamente por: MRAMONV el 10-07-2024 13:06:03
VELASQUEZ ESPINOZA BECKY KAZUMY <b>DNI:</b> 71403770 <b>ORCID:</b> 0000-0003-2481-6293	Firmado electrónicamente por: BVELASQUEZE el 10-07-2024 12:52:10

Código documento Trilce: INV - 1671356

## **Dedicatoria**

Dedicamos este trabajo a nuestras familias, cuyo apoyo incondicional ha sido nuestra fortaleza durante este viaje académico. A nuestros padres, quienes con su amor, paciencia y sabiduría nos enseñaron el significado de la perseverancia y dedicación.

## **Agradecimiento**

Queremos agradecer a nuestros asesores académicos por proporcionarnos orientación y conocimientos a lo largo de esta investigación. Su acompañamiento y respaldo han sido fundamentales, y sin ello, la realización de este proyecto no habría sido posible.

## Índice de contenidos

Carátula.....	i
Declaratoria de Autenticidad del Asesor.....	ii
Declaratoria de Originalidad de los Autores .....	iii
Dedicatoria .....	iv
Agradecimiento.....	v
Índice de tablas .....	vii
Índice de figuras .....	ix
Resumen .....	x
Abstract .....	11
I. INTRODUCCIÓN .....	2
II. METODOLOGÍA .....	17
2.1. Tipo y diseño de investigación .....	17
2.1.1. Tipo de investigación.....	17
2.1.2. Diseño de Investigación .....	17
2.2. Variables y operacionalización.....	18
2.3. Población (criterios de selección), muestra, muestreo, unidad de análisis. .	18
2.3.1. Población.....	18
2.4. Técnicas e instrumentos de recolección de datos .....	20
2.4.1. Técnica.....	20
2.4.2. Instrumento .....	20
2.4.3. Validez .....	21
2.5. Procedimientos .....	22
2.6. Métodos de análisis de datos.....	22
2.7. Aspectos éticos.....	23
III. RESULTADOS.....	24
II. DISCUSIÓN.....	47
III. CONCLUSIONES .....	50
IV. RECOMENDACIONES.....	52
REFERENCIAS.....	54
Anexos.....	62

## Índice de tablas

Tabla 1. Porcentaje de notas de la docente Maria Nina.....	13
Tabla 2. Porcentaje de notas de la docente Mariana Isabel Vitteri.....	14
Tabla 3. Resultados de juicio de experto: Validez de Instrumento.....	30
Tabla 4. Escala de calificaciones según el Minedu.....	31
Tabla 5. Frecuencia Estadístico de Rendimiento Académico 2021-2023.....	34
Tabla 6. Frecuencia estadística de la primera dimensión del 2023.....	35
Tabla 7. Frecuencia estadística de la segunda dimensión del 2023.....	35
Tabla 8. Frecuencia estadística de la tercera dimensión del 2023.....	36
Tabla 9. Frecuencia estadística de la primera dimensión del 2022.....	36
Tabla 10. Frecuencia estadística de la segunda dimensión del 2022.....	37
Tabla 11. Frecuencia estadística de la tercera dimensión del 2022.....	37
Tabla 12. Frecuencia estadística de la primera dimensión del 2021.....	38
Tabla 13. Frecuencia estadística de la segunda dimensión del 2021.....	38
Tabla 14. Frecuencia estadística de la tercera dimensión del 2021.....	39
Tabla 15. Medición de confiabilidad.....	41
Tabla 16. Estadística de fiabilidad.....	41
Tabla 17. Prueba de normalidad del 2021.....	42
Tabla 18. Prueba de normalidad del 2022.....	42
Tabla 19. Prueba de normalidad del 2023.....	43
Tabla 20. Escala para interpretar los valores de correlación de Spearman.....	44
Tabla 21. Prueba de Correlación de Spearman.....	45
Tabla 22. Prueba de Correlación de Spearman sobre la comunicación oral de la lengua materna.....	48
Tabla 23. Prueba de Correlación de Spearman sobre leer diversos tipos de textos escritos en su lengua materna.....	49
Tabla 24. Prueba de Correlación de Spearman sobre escribir diversos tipos de textos en su lengua materna.....	51
Tabla 25. Operacionalización de variables.....	65

Tabla 26. Requisitos funcionales y no funcionales.....	78
Tabla 27. Porcentaje de notas.....	81
Tabla 28. Porcentaje de Asistencias al año.....	82
Tabla 29. Porcentaje de participación al mes.....	82
Tabla 30. Porcentaje de asistencia de padres de familia a las reuniones por año.....	83

## Índice de figuras

Figura 1. Metodología CRISP-DM.....	14
Figura 2. Notas del área de comunicación de 5 y 6 grado de primaria del 2021 al 2023.....	26

## **Resumen**

La investigación realizada en la Institución Educativa Jorge Guevara Mellado evaluó el impacto de una aplicación web basada en árboles de decisión en el rendimiento escolar de 37 estudiantes de cuarto y quinto grado de primaria en el área de comunicación durante los años 2021-2023. Se utilizó un diseño no experimental transversal y un enfoque de investigación aplicada.

Los resultados mostraron que la mayoría de los estudiantes alcanzaron el logro esperado en las dimensiones de comunicación oral, lectura de textos y escritura en lengua materna. La prueba de confiabilidad indicó un valor de 0.904, evidenciando una buena consistencia en las calificaciones. Las pruebas de normalidad revelaron que los datos no se distribuían normalmente. Los valores de correlación de Spearman son 0.786, 0.65 y 0.47, indicando una relación fuerte entre las notas previas y las predichas por la aplicación.

En conclusión, la investigación demuestra que la aplicación web basada en árboles de decisión es una estrategia efectiva para mejorar el desempeño académico en comunicación, resaltando la importancia de integrar tecnologías educativas innovadoras en el entorno escolar.

**Palabras claves:** Rendimiento académico, aplicación web y árboles de decisión

## **Abstract**

The research carried out at the Jorge Guevara Mellado Educational Institution evaluated the impact of a web application based on decision trees on the academic performance of 37 fourth and fifth grade primary school students in the area of communication during the years 2021-2023. A non-experimental cross-sectional design and an applied research approach were used.

The results showed that most students achieved the expected achievement in the dimensions of oral communication, text reading and writing in the mother tongue. The reliability test indicated a value of 0.904, evidencing good consistency in the grades. Normality tests revealed that the data were not normally distributed. The Spearman correlation values are 0.786, 0.65 and 0.47, indicating a strong relationship between the previous grades and those predicted by the application.

In conclusion, the research demonstrates that the web application based on decision trees is an effective strategy to improve academic performance in communication, highlighting the importance of integrating innovative educational technologies in the school environment.

**Keywords:** Academic performance, web application and decision trees

## I. INTRODUCCIÓN

El reporte emitido por la Organización para la Cooperación y el Desarrollo Económicos (OCDE) en el año 2018 acerca de los resultados derivados del Programa Internacional de Evaluación de Estudiantes (PISA), destacaron las disparidades educativas a nivel global, evidenciando que los hallazgos revelan que el continente Asiático ostentó un liderazgo en cuanto al rendimiento en lectura y matemáticas, con más del 85% de los alumnos en nivel avanzados. En contraste, América Latina tuvo resultados por debajo del promedio en las tres áreas evaluadas, siendo México el único destacado con un 1% de los alumnos (Muelle, 2019, p.3). Estos hallazgos reflejan una situación inquietante: los niveles educativos en la región han permanecido constantes en las últimas dos décadas, lo que indica la persistencia de brechas educativas significativas. Esto subraya la urgente necesidad de abordar las desigualdades en el proceso educativo, implementando estrategias equitativas y efectivas para mejorar el rendimiento académico. La habilidad de expresarse verbalmente, comprender distintos tipos de textos y redactar diversos tipos de escritos en su idioma nativo es fundamental para la comprensión lectora, comunicación y expresión escrita. Según el OECD, 'El uso efectivo del lenguaje se posiciona como una herramienta fundamental para democracia y la implicación de la ciudadanía, y su uso efectivo es crucial para la implementación de políticas públicas y el servicio a los ciudadanos' (OECD, 2021, p. 12)

En este contexto, al examinar el desempeño estudiantil en el ámbito educativo, resulta imprescindible tener en cuenta una variedad extensa de factores que afectan su rendimiento. Aspectos como la situación socioeconómica familiar, los factores culturales y motivacionales, las destrezas personales y la participación familiar en el ámbito educativo son esenciales para el logro académico. (Ramirez-Vazquez, et al., p.20, 2020). La diversidad presente en los entornos familiares y culturales, junto con las variaciones en las habilidades y motivaciones de los

estudiantes, contribuye a la complejidad del proceso educativo. Además, la cooperación activa entre la institución educativa y los progenitores pueden ejercer una influencia notable en el rendimiento académico de los educandos. Esta interacción entre ambos entornos se transforma en un componente fundamental para abordar las brechas educativas y promover un rendimiento académico equitativo y sostenible.

En este escenario, la tesis se desarrolló en la Institución Educativa 20955-24 Jorge Guevara Mellado, ubicada en San Pedro de Mama, en el distrito de Santa Eulalia. Esta escuela, es un pilar en la comunidad, proporcionando educación a muchos niños. Sin embargo, a pesar de sus esfuerzos y dedicación, recientemente se han constatado falencias en las destrezas comunicativas y lingüísticas de los estudiantes de educación primaria, lo cual ha sido motivo de preocupación para la institución. Ante la crisis sanitaria desatada por la pandemia, estas carencias se vieron exacerbadas aún más, ya que muchos estudiantes enfrentan dificultades adicionales debido a la transición al aprendizaje remoto y la falta de acceso a recursos educativos adecuados. Esta situación exacerbó el deficiente desempeño académico del área de comunicación, lo que generó una mayor inquietud entre los docentes, la administración escolar y los propios familiares. Además, muchos niños durante la pandemia no tenían acceso a internet ni televisión, lo que limitaba aún más su capacidad para participar en el aprendizaje a distancia. Como evidencia del bajo rendimiento en la escuela, se han tomado las notas de los alumnos de los últimos tres años. Las figuras 1, 2, 3 y 4, que ilustran la situación, se encuentran en el anexo 9.

**Tabla 1. Porcentaje de notas de la docente Maria Nina**

Profesora	Año y grado	Nivel de logro	Porcentaje		
			01	02	03
			NL	NL	NL
	2023 - 5°	AD	18.75%	18.75%	0
		A	68.75%	50%	56.25%
		B	6.25%	31.25%	31.25%
		C	6.25%	0%	12.5%

Maria Nina	2022 – 4°	AD	18.75%	18.75%	0%
		A	68.75%	56.25%	56.25%
		B	6.25%	25%	31.25%
		C	6.25%	0%	12.5%
	2021 – 3°	AD	11.76%	0%	0%
		A	70.59%	58.82%	35.29%
		B	11.76%	41.18%	58.82%
		C	5.89%	0%	5.89%

Fuente: Realizado por el propio autor.

**Tabla 2.** *Porcentaje de notas de la docente Mariana Isabel Vitteri*

Profesora	Año y grado	Nivel de logro	Porcentaje		
			01	02	03
			NL	NL	NL
Mariana Isabel Vitteri	2023 - 4°	AD	16%	16%	0%
		A	64%	52%	60%
		B	12%	32%	32%
		C	8%	0%	8%
	2022 – 3°	AD	14.29%	9.52%	0%
		A	47.62%	57.14%	57.14%
		B	28.57%	33.33%	33.33%
		C	9.52%	0%	9.52%
	2021 – 2°	AD	9.52%	0%	0%
		A	52.38%	47.62%	47.62%
		B	33.33%	52.38%	52.38%
		C	4.76%	0%	0%

Fuente: Realizado por el propio autor.

A partir de esta circunstancia, se planteó a la Institución Educativa 20955-24 Jorge Guevara Mellado llevar a cabo una aplicación web que utiliza árboles de decisión con el fin de examinar la competencia académica de los estudiantes en el curso de comunicación. Esta herramienta analizará datos de los últimos tres años, incluyendo calificaciones y asistencia, para identificar tendencias y patrones en los desempeños estudiantiles en el ámbito educativo. El producto se mostrará visualmente en la página web de la institución, proporcionando al director información clave para tomar decisiones informadas sobre estrategias de mejora. Esta iniciativa aspira no sólo a elevar el rendimiento académico, sino también a

consolidar el compromiso de todos los involucrados y a propiciar un ambiente de aprendizaje cooperativo y dinámico. El objetivo es promover el éxito y el progreso académico integral de los educandos de la entidad educativa, preparándose para enfrentar desafíos académicos y personales a futuro.

A partir de las problemáticas detectadas en la investigación, se formuló la siguiente cuestión: ¿De qué manera se relaciona la aplicación web utilizando árboles de decisión con el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado?

Enseguida se presentaron los problemas específicos: ¿De qué manera se relaciona la aplicación web utilizando árboles de decisión con la comunicación oral de la lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado? ¿De qué manera se relaciona la aplicación web utilizando árboles de decisión con leer diversos tipos de textos escritos en su lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado? ¿De qué manera se relaciona la aplicación web utilizando árboles de decisión con escribir diversos tipos de textos en su lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado?

Justificación económica: La tecnología en las escuelas puede tener un impacto económico positivo a través del empoderamiento estudiantil y el mejoramiento del rendimiento académico. (Kalogiannidis, et al., p. 34, 2023). En tal sentido, el principal propósito de implementar esta herramienta tecnológica fue incrementar la rentabilidad académica de los estudiantes, lo que, a su vez, impulsó el progreso socioeconómico de la sociedad. Al optimizar la gestión de datos académicos, se buscó alcanzar una utilización más efectiva de los recursos financieros de la institución. Además, al anticipar el desempeño académico y disminuir la tasa de abandono escolar, se evitó la pérdida de recursos asociados a la formación de cada estudiante. Asimismo, el cumplimiento riguroso de normativas de privacidad y protección de datos fue una prioridad para evitar posibles multas o sanciones, lo

que contribuyó al ahorro de costos a largo plazo. De igual forma, cabe destacar que, si los estudiantes mejoren académicamente, se abrían oportunidades para que organizaciones donaran herramientas y recursos adicionales a la institución educativa, generando beneficios económicos adicionales.

Justificación social: Las tecnologías juegan un papel fundamental en promover la integración educativa y social en los centros escolares, al promover igualdad en el acceso al conocimiento y fomentar la participación inclusiva de todos los alumnos. (Reyes y Prado, p. 54. 2020). Su propósito principal fue brindar efectivas que promovieron el éxito educativo de los estudiantes. Al mejorar el rendimiento académico, se fortaleció el capital humano y se estimuló el crecimiento integral de los alumnos en todas sus facetas, teniendo un efecto positivo en su desarrollo personal y profesional. Además, este proyecto buscó fomentar la igualdad educativa al proporcionar recursos y posibilidades educativas para todos, sin importar su condición económica. En resumen, el estudio tuvo como objetivo crear un efecto positivo en la sociedad al elevar el nivel educativo y preparar a los estudiantes para la sociedad actual.

La justificación tecnológica: Incorporar herramientas tecnológicas en la educación facilita a los estudiantes el acceso a una variedad de información, promover la interactividad y el desarrollo de habilidades cognitivas, y prepararlos para el futuro con competencias digitales fundamentales. (Martínez-Garrido, p. 13, 2018). Esta herramienta tecnológica permitió un análisis detallado de los datos académicos de los estudiantes, identificando patrones y tendencias que ayuden a comprender mejor las necesidades individuales de cada alumno. Además, facilitó la detección anticipada de áreas que pueden mejorar, permitiendo intervenciones personalizadas y oportuna para los estudiantes en riesgo. La aplicación web también logró incrementar la eficacia en la administración de información, proporcionando una visión más clara y completa del desempeño estudiantil para informar decisiones y políticas educativas más efectivas. Por lo que, la aplicación web con árboles de decisión representa una herramienta tecnológica poderosa para abordar las deficiencias en el rendimiento académico y elevar el estándar

educativo en la institución escolar.

La justificación metodológica explicó las razones para seleccionar y emplear la metodología propuesta (Álvarez, 2020, p. 7). Por ende, se llevó a cabo el estudio mediante un diseño no experimental, integrando rúbricas como instrumentos de indagación. A través de un enfoque no experimental, se establecieron las metas de la investigación y se exploraron las conexiones entre causa y consecuencia, mientras que las rúbricas facilitaron la recolección de datos cualitativos y objetivos sobre el desempeño estudiantil. Asimismo, la elección de las rúbricas como instrumentos de evaluación se basó en su capacidad para ofrecer una evaluación detallada y objetiva del desempeño de los estudiantes, permitiendo una medición precisa de los criterios específicos de evaluación. Al utilizar rúbricas, se facilitó la estandarización de la evaluación y se promovió la consistencia en la calificación entre diferentes evaluadores. Estos enfoques garantizaron la precisión y confiabilidad de los resultados en el contexto escolar, al tiempo que facilitaron una comprensión más profunda de los elementos que afectan en el desempeño de los alumnos, aportando así al progreso del conocimiento en el contexto educativo.

Se cimentó el estudio en una fundamentación teórica que tuvo como objetivo identificar y abordar las áreas de conocimiento que necesitaban ser exploradas, respaldando la importancia del estudio con argumentos teóricos. Se optó por utilizar árboles de decisión como base teórica y método para analizar el rendimiento académico en la entidad educativa, en virtud de su significación y reconocimiento en el dominio de la computación cognitiva y el aprendizaje automatizado. Según Díaz, Ahumada y Melo (2021), los árboles de decisión son una técnica predictiva que busca evaluar los resultados de manera probabilística, facilitando la toma de decisiones óptimas a partir de observaciones específicas. (p. 12). El empleo de los árboles de decisión en la educación resultó crucial al posibilitar la identificación de factores y variables clave que ejercen influencia en la ejecución académica de los estudiantes. La eficacia de este enfoque metodológico para ahondar en la comprensión de las interacciones y

entrelazamientos que influyen en la excelencia del proceso de aprendizaje también fue resaltada por Timarán, Caicedo y Hidalgo (2018), quienes consideraron los árboles de decisión como el método predictivo más adecuado en la minería de datos para reconocer pautas vinculados al desempeño académico. (p. 6).

La incorporación de la teoría de árboles de decisión en el diseño de una aplicación web educativa tuvo como finalidad proporcionar una base robusta y científicamente validada para su desarrollo. Este fundamento teórico no solo justificó la selección de dicha metodología por su precisión y rigor metodológico, sino que también aseguró un enfoque sistemático para el procesamiento de información educativa. De esta manera, la teoría se transformó en una herramienta útil para tomar decisiones fundamentadas en el ámbito educativo, apoyando la creación de soluciones tecnológicas que respondieran a las necesidades de rendimiento académico.

El objetivo general fue: Determinar de qué manera se relaciona la aplicación web utilizando árboles de decisión con el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

A continuación, se detallaron los objetivos específicos: OE1: Determinar de qué manera se relaciona la aplicación web utilizando árboles de decisión en la comunicación oral de la lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado. OE2: Determinar de qué manera se relaciona la aplicación web utilizando árboles de decisión para leer diversos tipos de textos escritos en su lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado. OE3: Determinar de qué manera se relaciona la aplicación web utilizando árboles de decisión para escribir diversos tipos de textos en su lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

La hipótesis general planteada fue que: La aplicación web utilizando árboles de decisión influye significativamente con el rendimiento académico del curso de

comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

Las hipótesis específicas fueron las siguientes: HE1: La aplicación web utilizando árboles de decisión influye significativamente con la comunicación oral de la lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado. HE2: La aplicación web utilizando árboles de decisión influye significativamente con leer diversos tipos de textos escritos en su lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado. HE3: La aplicación web utilizando árboles de decisión influye significativamente al escribir diversos tipos de textos en su lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

Para desarrollar el marco teórico, se llevaron a cabo consultas en diversas bases de datos confiables, mayormente de nivel internacional. Un ejemplo de estudio relevante es el llevado a cabo por Salman (2017, p. 23), publicado en la revista Canadian Center for Science and Education, donde se examinó "El impacto de las aplicaciones web en el aula en la enseñanza, aprendizaje y rendimiento académico de las estudiantes de la facultad de educación". El objetivo central fue investigar cómo afecta el uso de aplicaciones web en la enseñanza y el rendimiento académico. En esta investigación experimental, se incluyeron 40 estudiantes elegidos de forma aleatoria. Los resultados mostraron que la integración de aplicaciones web mejora el entorno de aprendizaje y facilita las actividades educativas tanto para los docentes como para los estudiantes. Se destacó que el uso regular y efectivo de estas herramientas digitales, en combinación con una plataforma de gestión educativa, promueve la eficacia educativa.

Mendoza (2021) realizó una investigación llamada "Modelo de aplicación web 4.0 y su influencia en el rendimiento académico de estudiantes en la educación superior", publicado en la Revista Científica de Scielo. El objetivo principal fue desarrollar un modelo que pudiera proporcionar alertas tempranas basadas en

datos objetivos sobre mejorar el rendimiento académico y respaldar de forma eficaz el proceso de enseñanza, así como comprender de manera más precisa sus contextos individuales. Se realizó con una muestra de más de 400 estudiantes de la Universidad Privada Domingo Savio, utilizando un enfoque cuantitativo para el análisis de datos.

Los hallazgos alcanzados a través del uso de árboles de decisión en ciertas asignaturas permitieron discernir áreas susceptibles de mejoría en otras materias para los educandos involucrados. Específicamente, el modelo logró una precisión del 85% en el pronóstico del desempeño académico de los estudiantes y mejoró las calificaciones en un 12.5% en comparación con el rendimiento anterior. Además, el análisis de datos permitió identificar asignaturas específicas que necesitaban mejorar, lo que ayudó a los docentes a diseñar estrategias de enseñanza más efectivas. El modelo también proporciona alertas tempranas que posibilitaron que los profesores intervinieran de manera efectiva en el fomento de aprendizaje de los educandos, lo que llevó a un aumento significativo en el desempeño académico.

Rincón, et al. (2022), en el libro de la Conferencia Internacional sobre Interacción Humano-Computadora, investigaron el impacto del Monitoreo entre pares durante la pandemia mediante un sistema web para incrementar el desempeño académico de los estudiantes en el centro universitario colombiano. La investigación, la cual adoptó un enfoque cuantitativo con diseño descriptivo y experimental, evaluó a 463 estudiantes de diversas disciplinas en la Universidad CUC durante el año 2021. Los resultados sugieren que continuar con el seguimiento académico a través de la plataforma web sería beneficioso, ya que se observó una mejora positiva en el proceso educativo de los participantes.

A nivel nacional, se analizó el estudio de Aguilar (2020) sobre el empleo de Recursos REST y una aplicación web educativa para elevar el rendimiento académico en alumnos de primaria en la Institución Educativa de Cajamarca,

publicado en la UPN. El objetivo de este trabajo fue mejorar el desempeño en el ámbito de la cultura peruana, particularmente con discentes de 6° de la I.E.P "Nuestra Señora de Guadalupe". La investigación se enfocó en probar una hipótesis mediante un artefacto tecnológico, empleando un enfoque cuantitativo para evaluar su impacto en el desempeño académico. Los resultados mostraron una mejora significativa: mientras que inicialmente el 46.15% de los estudiantes tenía un bajo rendimiento, tras la implementación de la aplicación web, solo el 30.77% permaneció en esta categoría. Estos hallazgos respaldan la efectividad de la aplicación web para potenciar el desempeño académico y brindar acceso fluido a información relativa a las culturas del Perú.

Taza (2018) efectuó un estudio acerca de un tema llamado "Plataforma web para monitorear y gestionar el desempeño académico de estudiantes de secundaria en la institución educativa privada Jesús el maestro". El propósito del estudio fue examinar cómo un sistema informático podría encargarse de gestionar y supervisar el desempeño de los estudiantes de secundaria. Inicialmente, el estudio involucró a una muestra de 120 participantes, que incluía maestros, padres y personal académico. Luego, se seleccionó un grupo de 96 individuos para el análisis. El estudio se fundamentó en un enfoque cuantitativo. Los resultados demostraron altos niveles de acuerdo en términos de funcionalidad, confiabilidad, facilidad de uso y eficiencia (60.4%, 58.3%, 57.3%, 57.3% respectivamente).

Al revisar estudios previos relacionados, se sentaron las bases teóricas de nuestras variables para justificar y respaldar nuestra investigación.

Se tomó en cuenta como factor independiente la Aplicación Web, la cual se describe como un software que funciona en servidores remotos y se puede acceder a él a través de navegadores web, sin la necesidad de instalaciones locales. Esta cualidad proporciona a las aplicaciones web una versatilidad significativa, lo que permite su uso en diversos dispositivos y situaciones (Amapanqui, 2022).

En lo que respecta a la variable dependiente, se hace referencia al Rendimiento Académico. De acuerdo con Alboaneen, et al (2022), este aspecto es fundamental en el dominio educacional. Se enfatiza que los discentes con un rendimiento deficiente enfrentan diversos desafíos, por lo que es crucial que las instituciones educativas supervisen de cerca su progreso y ofrezcan apoyo temprano a aquellos que lo necesiten (p.8). Según el MINEDU, el rendimiento académico se define como el grado en que un estudiante ha adquirido y puede demostrar una serie de habilidades, conocimientos y actitudes, lo cual se refleja en un sistema de evaluación y calificaciones que indican su situación educativa actual. Según lo establecido en el Currículo Nacional de Educación Básica, la evaluación del desempeño académico se fundamenta en la evaluación de las competencias, habilidades y niveles de aprendizaje de cada uno de los estudiantes han logrado al concluir su ciclo educativo. Dichas competencias, que comprenden una gama de conocimientos, habilidades y actitudes, constituyen los elementos que los estudiantes aplican para actuar de forma efectiva ante retos particulares. Por su parte, el desempeño académico se refiere a la manifestación concreta y mensurable del avance en el desarrollo de estas competencias, visible en una variedad de escenarios educativos y sociales, proporcionando un reflejo más amplio de la capacidad del estudiante para aplicar su aprendizaje en la práctica. Según Ariza, et al (2018), en el contexto educativo, el rendimiento académico se concibe como la relación entre los recursos invertidos (tales como personalidad, motivación, habilidades e intereses, entre otros) y los logros alcanzados. Esto implica que el rendimiento académico está afectado por aspectos como las capacidades cognitivas, la personalidad, los métodos de estudio, la autoestima, entre otros factores. Por consiguiente, lo consideran como un indicador de la calidad del aprendizaje del estudiante. La ejecución académica de los estudiantes es significativamente influenciada conforme al Currículo Nacional de Educación Básica, el cual establece los objetivos y contenidos fundamentales del aprendizaje. Este currículo sirve como una guía para la planificación educativa y el análisis del

aprendizaje.

La valoración del desempeño académico se lleva a cabo mediante el análisis de competencias, destrezas y niveles de aprendizaje alcanzados al concluir la educación (MINEDU, 2019). En el área de comunicación, el MINEDU pone énfasis en el desarrollo de habilidades comunicativas mediante el uso del lenguaje para interactuar, entender la realidad y representar el mundo. Estas habilidades comprenden la lectura, comunicación oral y escritura, con el objetivo de preparar a los estudiantes para comunicarse eficazmente en diversas situaciones y reflexionar sobre los mensajes que reciben y generan.

En cuanto a la primera dimensión, que también es una de las competencias definidas por el MINEDU y se titula "se comunica oralmente en su lengua materna", esta se define como: La competencia de comunicación oral en la lengua materna implica expresar y entender ideas al interactuar con otros. Esta habilidad social fomenta el uso creativo y responsable del lenguaje, promoviendo una visión crítica de los medios audiovisuales, lo que ayuda a formar identidades y a crecer personalmente. De lo cual, implica la combinación de 6 capacidades, en esta ocasión para el estudio se consideró solamente 3 (p.40).

Según el MINEDU y en relación con la segunda dimensión, la competencia de "lectura de diversos tipos de textos escritos en lengua materna" implica que el estudiante no solo entiende la información, sino que también la interpreta y reflexiona sobre los textos, utilizando su experiencia y contexto sociocultural. Esta capacidad promueve una comprensión crítica, que es fundamental en un mundo influenciado por las nuevas tecnologías, y contribuye al desarrollo personal y social del estudiante (p. 44).

Según el MINEDU y en relación con la tercera dimensión, la competencia de "redacción de diversos tipos de textos en lengua materna" implica que el estudiante utilice el lenguaje escrito para construir y comunicar significados. Este

proceso reflexivo implica la organización y revisión de textos de acuerdo con el contexto y el propósito, haciendo uso de su experiencia y diversas estrategias. Esta habilidad fomenta una comunicación escrita creativa y responsable, la cual resulta esencial para la interacción social y el crecimiento individual. (p. 46).

Siguiendo con los fundamentos conceptuales, se presentan herramientas y conceptos fundamentales de tecnología. En este sentido, según Contreras, et al. (2020) definen al Machine Learning en relación con la educación como el proceso de recopilar, analizar y usar datos, con la finalidad de comprender y perfeccionar el proceso de adquisición de conocimientos. Además, los autores señalan que Machine Learning cuenta con diferentes algoritmos para predecir resultados, como regresión logística, árboles de decisión, K-Means, SVM y entre otros.

En una investigación realizada por Gutiérrez, et al. (2022), se exploró el pronóstico del desempeño académico de los estudiantes mediante técnicas de Aprendizaje Automático. Al evaluar técnicas como árboles de decisión, bosques aleatorios y redes neuronales, encontraron que los árboles de decisión tuvieron una mayor precisión en la predicción, lo que puede ser útil para establecer políticas de apoyo a alumnos en riesgo académico.

En el marco de esta situación, los árboles de decisión se destacan como una herramienta valiosa y efectiva para evaluar el rendimiento académico. Estos árboles, que representan una técnica de modelado y visualización, emplean reglas lógicas y datos anteriores para tomar decisiones. Su uso facilita la identificación de factores clave, la predicción de resultados y el desarrollo de recomendaciones personalizadas acorde a las necesidades individuales de los alumnos. Los DT se caracterizan por ser fáciles de entender y por proporcionar conclusiones precisas, lo que los hace accesibles tanto para profesionales de la educación como para personas sin experiencia especializada en el campo. Además, su versatilidad y capacidad para integrarse con otras técnicas de aprendizaje automático lo convierte en un recurso adaptable para potenciar el desempeño académico.

La construcción del aplicativo web se fundamentó en la elección de Python como

el lenguaje de programación central, una selección estratégica debido a su sintaxis clara e intuitiva que promueve la agilidad en la evolución del proyecto y facilita un mantenimiento eficiente del código a largo plazo.

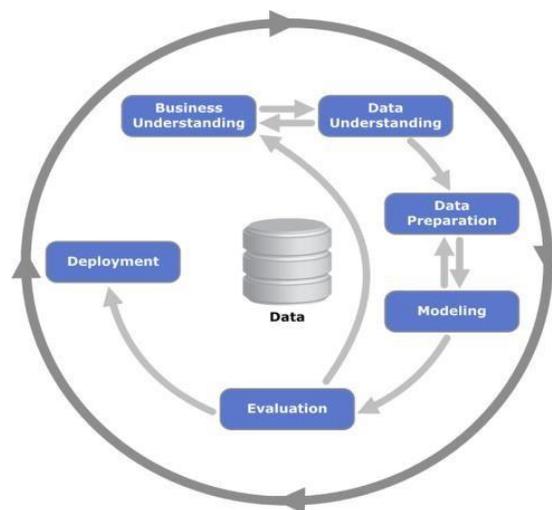
Python se distingue por su alto nivel de abstracción y su naturaleza versátil, lo que permite a los desarrolladores transitar por una curva de aprendizaje razonable y aplicarlo en un espectro amplio de situaciones. La simplicidad de su estructura gramatical fomenta un estilo de escritura directo y expresivo, conduciendo a la creación de código no solo eficiente sino también estructurado y comprensible, un activo valioso para el trabajo colaborativo y la cooperación entre equipos multidisciplinarios. Además, Python se beneficia de un sistema de tipado dinámico que ofrece flexibilidad en la definición de tipos de datos, y su condición de lenguaje interpretado facilita las pruebas y la iteración en tiempo real. Estas características lo convierten en una solución idónea para el prototipado rápido y para ciclos de desarrollo y despliegue ágiles en distintas áreas, incluyendo el desarrollo web y el análisis de datos.

Soto (2021) resalta la notable adaptabilidad de Python, enfatizando su accesibilidad y amplio rango de aplicaciones prácticas. Su capacidad para operar de manera efectiva en diversas plataformas y su compatibilidad con otras tecnologías lo hacen excepcionalmente adecuado para proyectos de gran envergadura que requieren escalabilidad y una integración fluida con variados sistemas y herramientas. (p. 13).

En concordancia con estas apreciaciones, Chuquimarca, Pinzón y Rosales (2021) reconocen a Python como un lenguaje de programación de alto nivel e interpretado, operando bajo un modelo en el cual el código es ejecutado directamente, obviando la compilación. Destaca su habilidad para manejar diversos paradigmas de programación, incluyendo la orientación a objetos, la programación imperativa y la funcional, lo que resalta su versatilidad y amplio rango de uso. (p. 23)

En esta investigación, se utilizó la metodología CRISP-DM (Proceso Estándar Cruzado para la Minería de Datos) como estructura para realizar el análisis de información. CRISP-DM se compone de seis etapas esenciales: comprensión del contexto empresarial, recopilación de información, tratamiento de datos, modelización, valoración e implementación (Schröe, Kruse y Gómez, 2021, p. 13). La primera etapa establece los requisitos y objetivos del proyecto desde una óptica empresarial. En la segunda etapa, se recopilan y se familiarizan con los datos disponibles. La tercera etapa se centra en limpiar y transformar los datos para prepararlos para el análisis. Durante la cuarta etapa, se aplican técnicas de modelado, como los árboles de decisión, para identificar patrones y relaciones. La quinta etapa evalúa la calidad y la eficacia de los modelos desarrollados. Finalmente, en la sexta etapa, los resultados se integran en el entorno operativo para luego usarse en la toma de decisiones.

Esta metodología estructurada asegura un enfoque sistemático y eficaz para convertir los datos educativos en información valiosa y aplicable (Durando, Giraldo Vargas y Soto, 2023, p. 8).



**Figura 1.** Metodología CRISP-DM. Extraído por: IPMOGUIDE

## II. METODOLOGÍA

### 2.1. Tipo y diseño de investigación

#### 2.1.1. Tipo de investigación

El método de estudio empleado se diferenció por ser aplicada, ya que se centró en abordar un problema específico a través de la investigación y su posterior implementación. Esta elección se justificó por la necesidad de encontrar soluciones prácticas y concretas para mejorar las estrategias actuales.

La investigación aplicada se destacó por analizar situaciones sociales específicas y proponer medidas concretas para mejorar las estrategias, desarrollarlas y aplicarlas en respuesta a las necesidades identificadas (Castro, et al. 2023, p.15). Además, se basó en los hallazgos obtenidos de la investigación básica en ciencias naturales y sociales. A partir de estos resultados, se detectaron problemáticas y se elaboraron hipótesis de trabajo para abordar los retos de la vida social en una comunidad a escala regional o nacional. (Ñaupas et al. 2018, p.136).

#### 2.1.2. Diseño de Investigación

Se empleó un diseño de investigación no experimental de tipo transversal, dado que no se manipuló la variable independiente y se observaron las relaciones entre variables en su entorno natural, lo que posibilitó estudiar el rendimiento académico sin intervenir directamente en él.

Según Hernández y colaboradores (2018), la investigación no experimental se categoriza según la temporalidad o la cantidad de momentos en los que se recopilan los datos. En ciertas ocasiones, se concentra en analizar variables, evaluar situaciones o determinar relaciones en un único momento, empleando un diseño transversal. En otros casos, se examinan las variables y sus cambios a lo largo del tiempo, utilizando un diseño longitudinal (p.154). Estos mismos

autores indican que los diseños de investigación transversales implican el análisis de datos de distintos grupos en un único momento sin intervenir en las variables, observando las relaciones entre variables tal como ocurren en la realidad (p.186).

Además, se utilizó un enfoque de estudio correlacional, tiene como fin estudiar la interdependencia entre dos o más variables. En la investigación no se busca establecer relaciones causales, sino determinar y cuantificar la intensidad y dirección de las asociaciones entre las variables estudiadas. La investigación correlativa es particularmente beneficiosa en estudios no experimentales, ya que permite entender cómo interactúan diversos factores en su entorno natural sin necesidad de alterar variables.

## 2.2. Variables y operacionalización

**Variable dependiente:** Rendimiento Académico

**Definición conceptual:** El rendimiento académico se refiere al conocimiento que los estudiantes adquieren como consecuencia de la enseñanza del profesor. (Minedu, 2015, p.13)

**Definición Operacional:** Según el Ministerio de Educación (MINEDU) el desempeño académico del curso de comunicación puede medirse a través de indicadores organizados en tres dimensiones (2019, p.164).

**Indicador:** El criterio escogido para evaluar la variable dependiente es: Rendimiento Académico.

Se utilizó una escala ordinal para medir la variable dependiente.

La descripción detallada de la matriz operacionalización de la variable está ubicado en el Anexo 1.

## 2.3. Población (criterios de selección), muestra, muestreo, unidad de análisis.

### 2.3.1. Población

En el contexto del estudio, todos los individuos o elementos involucrados en el análisis constituyen el conjunto íntegro de elementos que constituyen la

población del objeto de estudio (Condori y Ojeda, 2020, p. 15). Según

Satishprakash (2020), todos los elementos o unidad de análisis que estuvieron disponibles en el ámbito específico de la indagación forman parte de la población. (p. 8). Por consiguiente, el enfoque de la investigación se concentró en analizar a estos individuos para establecer una base sólida y obtener resultados representativos de la población objetivo. El grupo de estudiantes que participó en la investigación estuvo compuesto por 42 alumnos de los grados de quinto y sexto de primaria.

Criterios de selección: Estudiantes matriculados en el grado de quinto y sexto de primaria.

Criterios para no incluir: Estudiantes que no estuvieron matriculados en la institución en los tres últimos años y no cuentan con registros completos de calificaciones y asistencia

La muestra consiste en un grupo más reducido que se considera representativo de toda la población en estudio. Los datos se obtienen de esta muestra, delimitando así la población según el problema de investigación. Según Hernandez, Sampieri y Mendoza (2018), una muestra proporciona información sobre una parte significativa de la población. (p. 16) Dado que la población de interés era inferior a 100 individuos, no fue necesario calcular una muestra representativa. Se evaluaron únicamente a 37 estudiantes de quinto y sexto grado de primaria de la institución educativa.

Se empleó un método de muestreo intencional o conveniente. En este enfoque, el investigador elige los elementos de la muestra sin recurrir a la probabilidad, seleccionando aquellos que cumplen con ciertas características para lograr mejores resultados. Según Sucasaire (2022), este método es común en investigaciones cualitativas, estudios piloto o exploratorios. (p. 6). Ruiz y Valenzuela (2018) señalan que el muestreo no aleatorio puede ser adecuado para obtener una muestra que se asemeje adecuadamente a una población,

yasea limitada o ilimitada, con el propósito de calcular parámetros estadísticos.  
(p.

14). La correcta definición de la muestra puede ser crucial para influir en los resultados tanto en la práctica como en la teoría (Cash et al., 2022, p.15).

## 2.4. Técnicas e instrumentos de recolección de datos

### 2.4.1. Técnica

Se optó por el método de análisis documental para la recolección de información. El análisis documental se usó para reunir las calificaciones históricas y actuales, junto con los registros de asistencia de los estudiantes.

Tabla 3. Resultado de juicio de experto: Validez de instrumento

Experto	Grado	Juicio
Mgtr. Aurón Mendoza Canales	Magister	Aplicable
Dr. Giancarlo Sanchez Atuncar	Doctor	Aplicable
Mgr. Dany Montoya Negrillo	Magister	Aplicable

Fuente: Elaboración propia.

### 2.4.2. Instrumento

El método empleado para recopilar información consistió en la rúbrica de calificación o de evaluación de comunicación, ya que es empleada por los profesores para valorar o medir la competencia académica de los alumnos, estableciendo criterios específicos para las calificaciones de AD (Nivel Destacado), A (Nivel Esperado), B (En proceso) y C (Inicio).

**Tabla 4.** *Escala de calificaciones según el Minedu.*

AD	Logro Destacado	Demuestra un aprendizaje que va más allá de lo esperado cuando el estudiante supera las expectativas en cuanto a la competencia.
A	Logro Esperado	En el tiempo establecido y en todas las tareas propuestas, el estudiante muestra un manejo satisfactorio cuando alcanza el nivel esperado en la competencia.
B	En Proceso	Cuando el estudiante se acerca al nivel esperado por la competencia, necesita recibir acompañamiento durante un periodo razonable para alcanzarlo.
C	En Inicio	Cuando el estudiante muestra problemas recurrentes en la realización de tareas y presenta avances mínimos en una habilidad esperada, requiere mayor apoyo y atención del profesor.

Según García (2019) indica que la creación de una rúbrica implica fijar metas, identificar los elementos a evaluar para garantizar la objetividad, definir descriptores y escalas de calificación, asignar pesos a cada criterio, y revisar la rúbrica elaborada para reflexionar sobre su impacto educativo, resaltando su utilidad como herramienta en la labor docente. (p.19)

#### 2.4.3. Validez

Es un instrumento de medición que cuantifica de manera significativa y adecuada el fenómeno que se desea evaluar.

Según Urrutia Egaña et al. (2014) mencionan tres tipos principales de validez:

- Validez de contenido: Se evalúa según la adecuación de los ítems o preguntas para cubrir de manera representativa el área de estudio y el constructo a evaluar.
- Validez de constructo: Se trata de la exactitud con la que un instrumento

mide el concepto teórico o constructo que intenta medir. Se analiza si las puntuaciones del instrumento están relacionadas de manera consistente con otras variables (Ventura-León, 2017, p. 12).

- Validez predictiva: Indica la capacidad del instrumento para predecir un criterio o variable de interés en el futuro. Se establece mediante la correlación entre las puntuaciones del instrumento y los valores de la variable criterio (Hernández-Nieto, 2018, p. 14).

Con el propósito de asegurar la calidad de los resultados de la investigación, es esencial que los instrumentos sean válidos y confiables. Esto implica evaluar la fiabilidad del instrumento utilizado el coeficiente alfa de Cronbach, que indica la consistencia interna de las mediciones. (Toro-Huamanchumo, et al., 2019, p.15)

## 2.5. Procedimientos

Los registros históricos y/o actuales de calificaciones y la asistencia de los estudiantes fueron recopilados mediante análisis documental y depositados en una base de datos para un análisis posterior. Una vez recolectados, estos datos fueron ingresados al software estadístico SPSS v22 para su procesamiento y análisis. Se realizaron pruebas de normalidad utilizando Shapiro-Wilk, dada la muestra de tamaño reducido. Además, se exploraron las relaciones entre el rendimiento académico y otros factores relevantes mediante técnicas estadísticas descriptivas e inferenciales.

Se utilizaron gráficos de barras y tablas para presentar los resultados de manera visual, asegurando una interpretación clara y una visualización efectiva. Estos gráficos mostraron el desempeño académico en las diversas dimensiones evaluadas.

## 2.6. Métodos de análisis de datos

El procesamiento de datos emerge como una etapa fundamental en el proceso investigativo, ya que implica la organización de la información recopilada para realizar un análisis exhaustivo. En esta etapa, los datos se describen, caracterizan

e interpretan meticulosamente. El tipo de análisis ya sea cualitativo, cuantitativo o una combinación de ambos, depende de los objetivos y necesidades específicas del estudio. (Sánchez, Reyes y Mejía, 2018, p. 26).

Se utilizaron métodos de estadísticas descriptivas para resumir y organizar los datos recopilados, incluyendo el cálculo del promedio del rendimiento académico en comunicación y la desviación estándar para comprender la variabilidad en dicho rendimiento. Además, se emplearon técnicas de estadística inferencial para investigar hipótesis y obtener conclusiones más generales sobre la población estudiada. Se comprobó la normalidad de los datos utilizando la prueba de Shapiro-Wilk, ya que la muestra era menor a 50 (Molina, 2022, p.7), y se establece un nivel de significancia del 5% ( $p < 0.05$ ) para determinar la relevancia estadística de los resultados.

## 2.7. Aspectos éticos

Se observaron las directrices establecidas por las normas ISO y los principios fundamentales de investigación, en consonancia con el código ético de la institución educativa. Esto aseguró la integridad y precisión de los datos recopilados.

### III. RESULTADOS

Se aplicó un método de investigación aplicada con un diseño no experimental de tipotransversal para examinar cómo se desempeñaron académicamente los 37 alumnos

de los dos grados de la escuela primaria. Por lo cual, se recopiló información histórica y actual de calificaciones y registros de asistencia mediante análisis documental y unarúbrica de calificación. Los datos fueron procesados con el software SPSS v22, empleando técnicas de estadística descriptiva e inferencial para analizar el rendimiento en el área de comunicación y proponer estrategias educativas más efectivas.

Se utilizó el análisis descriptivo para sintetizar las características clave de los datos, como el promedio de rendimiento académico, la desviación estándar y otras medidas relevantes. Esto ofreció una visión general de cómo estaban distribuidas las calificaciones y de su variabilidad. Los hallazgos permitieron obtener una mayor comprensión de cómo se distribuyen las calificaciones y la consistencia del rendimiento académico en el grupo estudiado. Se creó una tabla de frecuencias para resumir la distribución de las calificaciones en comunicación durante el año académico 2023, identificando patrones o tendencias en las calificaciones obtenidas.

**Tabla 5. Frecuencia Estadístico de Rendimiento Académico 2021-2023**

	D1 - 2021	D2 - 202 1	D3 - 202 1	D1 - 202 2	D2 - 202 2	D3 - 202 2	D1 - 202 3	D2 - 202 3	D3 - 202 3
N Válido	37	37	37	37	37	37	37	37	37
Perdidos	0	0	0	0	0	0	0	0	0
Media	2,756 8	2,54 05	2,40 54	2,81 08	2,81 08	2,45 95	2,89 19	2,81 08	2,45 95
<b>Error estándar de la media</b>	,1188 4	,083 06	,090 54	,133 32	,108 48	,113 60	,127 25	,108 48	,113 60
<b>Mediana</b>	3,000 0	3,00 00	2,00 00	3,00 00	3,00 00	3,00 00	3,00 00	3,00 00	3,00 00
<b>Desviación estándar</b>	,7228 6	,505 23	,550 73	,810 96	,659 88	,691 00	,774 01	,659 88	,691 00

Varianza	,523	,255	,303	,658	,435	,477	,599	,435	,477
Asimetría	-524	-	-	-	,219	-	-	,219	-
Error estándar de asimetría	,388	,170	,131	,623	,915	,947	,219	,915	,388
Mínimo	1,00	2,00	1,00	1,00	2,00	1,00	1,00	2,00	1,00
Máximo	4,00	3,00	3,00	4,00	4,00	3,00	4,00	4,00	3,00

Fuente: Resultados generados con SPSS V22, elaboración propia.

Interpretación: La tabla 5 revela una distribución asimétrica con la mayoría de los estudiantes obteniendo calificaciones cercanas a 3. Tanto la media como la mediana de las calificaciones por periodo son similares, lo que indica consistencia en el rendimiento académico a lo largo del tiempo. La desviación estándar y el rango de calificaciones muestran cierta variabilidad en el rendimiento, con algunos estudiantes obteniendo calificaciones altas o bajas. Se presenta en la tabla 6 una síntesis detallada de cómo están distribuidas las calificaciones y la variabilidad en el rendimiento académico, lo que ayuda a identificar patrones y tendencias, y facilita la creación de estrategias educativas más eficaces.

**Tabla 6.** Frecuencia estadística de la primera dimensión del 2023

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido En inicio (C)	3	8,1	8,1	8,1
En proceso (B)	4	10,8	10,8	18,9
Logro esperado (A)	24	64,9	64,9	83,8
Logro destacado (AD)	6	16,2	16,2	100,0
Total	37	100,0	100,0	

Fuente: Resultados generados con SPSS V22, elaborada por el autor.

Análisis: A partir de los resultados presentados en la tabla 6, se observa que el 64,9% de los alumnos alcanzaron el nivel esperado (A), mientras que el 16,2% logró un nivel destacado (AD). Además, un 10,8% está en proceso (B) y solo un 8,1% se encuentra en nivel inicial (C). Esto indica que hay un grupo considerable que podría beneficiarse de mejoras en su capacidad para comunicarse oralmente.

**Tabla 7. Frecuencia estadística de la segunda dimensión del 2023**

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	En proceso (B)	12	32,4	32,4
	Logro esperado (A)	20	54,1	86,5
	Logro destacado (AD)	5	13,5	100,0
	Total	37	100,0	100,0

Fuente: Resultados generados con SPSS V22, elaborada por el autor.

Análisis: Según se desprende de los datos presentados en la tabla 7 los estudiantes (54,1%) alcanzaron el nivel esperado (A), un 32,4% se encuentran en proceso (B), y un 13,5% lograron un nivel destacado (AD). Esto sugiere que la mayoría de los estudiantes cumplieron con las expectativas, mientras que una minoría superó dichas expectativas.

**Tabla 8. Frecuencia estadística de la tercera dimensión del 2023**

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	En inicio (C)	4	10,8	10,8
	En proceso (B)	12	32,4	43,2
	Logro esperado (A)	21	56,8	100,0
	Total	37	100,0	100,0

Fuente: Resultados generados con SPSS V22, elaborada por el autor.

Análisis: Se visualiza en la tabla 8 que en la mayoría de los estudiantes (56,8%) lograron el "Logro esperado" (A), un 32,4% están en proceso (B), y un 10,8% se encuentran en la fase inicial (C). Esto indica que la mayoría ha alcanzado los estándares esperados, aunque algunos aún están en proceso de desarrollo.

Este mismo procedimiento de análisis será replicado para los años 2022 y 2021, con el propósito de contrastar el rendimiento académico a lo largo de los años y discernir posibles tendencias o mejoras en el desempeño de los alumnos. La tabla siguiente muestra la distribución de calificaciones de comunicación de los 37 alumnos durante el año 2022.

**Tabla 9.** Frecuencia estadística de la primera dimensión del 2022

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	En inicio (C)	3	8,1	8,1	8,1
	En proceso (B)	7	18,9	18,9	27,0
	Logro esperado (A)	21	56,8	56,8	83,8
	Logro destacado (AD)	6	16,2	16,2	100,0
	Total	37	100,0	100,0	

Fuente: Resultados generados con SPSS V22, elaborada por el autor.

Análisis: Se revela en la tabla 9 que, en el año 2022, un 56.8% de los estudiantes alcanzaron el "Logro esperado" (A) en comunicación oral en su lengua materna, mientras que un 16.2% lograron un "Logro destacado" (AD). Un 18.9% de los estudiantes están en proceso (B), y un 8.1% se encuentran en la fase inicial (C). Estos datos indican un buen desempeño general en la competencia de comunicación oral, con la mayoría de los estudiantes cumpliendo o superando las expectativas, aunque algunos todavía están en desarrollo.

**Tabla 10.** Frecuencia estadística de la segunda dimensión del 2022

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	En proceso (B)	12	32,4	32,4	32,4
	Logro esperado (A)	20	54,1	54,1	86,5
	Logro destacado (AD)	5	13,5	13,5	100,0
	Total	37	100,0	100,0	

Fuente: Resultados generados con SPSS V22, elaboración propia.

Interpretación: En el año 2022, según la tabla 10, la gran parte de los discentes lograron el nivel esperado de competencia en "Leer diversos tipos de textos escritos en su lengua materna", con un 54.1% alcanzados el "Logro esperado" y un 13.5% obteniendo un "Logro destacado". Un 32.4% de los estudiantes se encontraban en procesos de alcanzar estos niveles. Estos hallazgos reflejan que la amplia mayoría de los estudiantes cumplieron con el nivel esperado, mientras que una minoría superó las expectativas establecidas.

**Tabla 11. Frecuencia estadística de la tercera dimensión del 2022**

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	En inicio (C)	4	10,8	10,8	10,8
	En proceso (B)	12	32,4	32,4	43,2
	Logro esperado (A)	21	56,8	56,8	100,0
	Total	37	100,0	100,0	

Fuente: Resultados generados con SPSS V22, elaboración propia.

Análisis: Los resultados de la tabla 11, muestran que un 56.8% de los estudiantes lograron el nivel de “Logro esperado” en la capacidad de “Escribir diversos tipos de textos en su lengua materna”. Un 32,4% están en proceso de alcanzar este nivel, mientras que un 10.8% se encuentran en la fase inicial. Estos datos sugieren que la gran parte de los discentes demuestran un desempeño adecuado en la escritura, aunque algunos aún necesitan mejorar.

También se elaboró una tabla de frecuencias para cada dimensión del rendimiento académico en el año 2021.

**Tabla 12. Frecuencia estadística de la primera dimensión del 2021**

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	En inicio (C)	2	5,4	5,4	5,4
	En proceso (B)	9	24,3	24,3	29,7
	Logro esperado (A)	22	59,5	59,5	89,2
	Logro destacado (AD)	4	10,8	10,8	100,0
	Total	37	100,0	100,0	

Fuente: Resultados generados con SPSS V22, elaborada por el autor.

Análisis: La gran parte de los discentes representando el 59.5% según la tabla 12, alcanzaron el nivel de logro esperado en comunicación oral en su lengua materna, mientras que un 10.8% lograron un logro destacado. Un 24.3% están en proceso de mejorar y un 5.4% están en la fase inicial. Esto muestra que la mayoría de los estudiantes han alcanzado el nivel esperado en habilidades de comunicación oral, aunque algunos aún tienen un margen para mejorar.

**Tabla 13. Frecuencia estadística de la segunda dimensión del 2021**

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido En proceso (B)	17	45,9	45,9	45,9
Logro esperado(A)	20	54,1	54,1	100,0
Total	37	100,0	100,0	

Fuente: Resultados generados con SPSS V22, elaboración propia.

Análisis: De acuerdo con la tabla 13, la gran parte de los estudiantes, un 54.1%, alcanzaron el "Logro esperado" (A) en la competencia de "Leer diversos tipos de textos escritos en su lengua materna", mientras que un 45.9% se encuentran "En proceso" (B). Estos descubrimientos apuntan a que la preponderancia de estudiantes ha alcanzado el nivel esperado en la comprensión de textos escritos en su lengua materna, aunque hay una proporción considerable que aún necesita mejorar su desempeño.

**Tabla 14. Frecuencia estadística de la tercera dimensión del 2021**

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido En inicio (C)	1	2,7	2,7	2,7
En proceso (B)	20	54,1	54,1	56,8
Logro esperado (A)	16	43,2	43,2	100,0
Total	37	100,0	100,0	

Fuente: Resultados generados con SPSS V22, elaboración propia.

Análisis: Según los datos presentados en la tabla 14, la gran parte de los alumnos, un 54.1%, se ubican "En proceso" (B) de alcanzar el estándar deseado en la competencia de "Escribir diversos tipos de textos en su lengua materna". Un 43.2% de los estudiantes han logrado el "Logro esperado" (A), mientras que un 2.7% se encuentran en la fase inicial (C). Estos resultados indican que, aunque una proporción significativa de estudiantes está en proceso de mejorar su rendimiento en la escritura, una parte considerable ya ha alcanzado el nivel esperado. Para evaluar las destrezas de comunicación oral y escrita en los grados quinto y sexto,

se han recopilado las calificaciones en estas áreas durante tres años consecutivos con el fin de analizar las tendencias y el progreso de los estudiantes.



**Figura 2.** Notas del área de comunicación de 5 y 6 grado de primaria del 2021-2023.

Interpretación: El gráfico muestra la distribución de las notas por categoría (A, B, C) y por año (2021, 2022, 2023). En general, se observa que la mayoría de las notas son de tipo A, seguidas por las de tipo B y en menor cantidad las de tipo C. Esta tendencia se mantiene constante a lo largo de los tres años.

#### Prueba de confiabilidad

En el transcurso de esta investigación, se realizó una evaluación exhaustiva de la confiabilidad de las calificaciones registradas por 37 estudiantes durante tres años académicos. La necesidad de esta evaluación radica en la importancia de asegurar que las calificaciones reflejen de manera precisa y consistente el desempeño de los alumnos. Dado que cada estudiante tiene tres calificaciones por año, se empleó el coeficiente alfa de Cronbach como un indicador de consistencia interna y fiabilidad en el análisis de datos, el cual constituye una medida ampliamente aceptada para evaluarla consistencia interna de un conjunto de ítems, como las calificaciones, particularmente útil cuando se busca medir un constructo complejo, como el rendimiento académico.

La aplicación del alfa de Cronbach a las nueve calificaciones (tres por año)

permitió determinar si las notas están en línea con las expectativas y el desempeño real de los estudiantes a lo largo del tiempo. Este enfoque no solo asegura la consistencia de las calificaciones, además de ello, ofrece una base robusta para la interpretación de los resultados. Es crucial resaltar que el alfa de Cronbach simplifica el proceso de evaluación al requerir solo una administración del instrumento de evaluación. Además, este coeficiente genera un valor entre 0 y 1, donde valores más altos indican una mayor consistencia en las calificaciones. En este sentido, un valor alto de alfa de Cronbach respalda la fiabilidad de las calificaciones y, por ende, la validez de las conclusiones derivadas de este estudio (Tuapanta, Duque y Mena, 2017, p. 5).

**Tabla 15. Medición de confiabilidad**

Valor	Apreciación
0 a 0.49	No fiable (inaceptable)
0.50 a 0.59	No fiable (deficiente)
0.60 a 0.69	Fiable en menor medida (cuestionable)
0.70 a 0.75	Fiable aceptable
0.76 a 0.89	Fiable considerable (buena)
0.90 a 1	Fiable altamente (excelente)

Fuente: George, y Mallery (2003). Criterios de confiabilidad. p. 231.

Se analizó la coherencia de las calificaciones de los estudiantes durante los últimos tres años mediante el cálculo del coeficiente alfa de Cronbach utilizando el Software SPSS V22.

**Tabla 16. Estadística de fiabilidad**

Alfa de Cronbach	N de elementos
,904	9

Fuente: Resultados obtenidos mediante el SPSS V22, desarrollo propio.

Interpretación: Según la tabla 16, se alcanzó un coeficiente alfa de Cronbach de 0.904 para los nueve ítems evaluados, lo que muestra una fuerte coherencia interna entre estos ítems. Esto sugiere que las mediciones realizadas son confiables y consistentes.

### Prueba de Normalidad

Después de verificar la fiabilidad de los datos, se realizó un análisis para comprobar si estos siguen una distribución normal. Este paso reviste una importancia crucial, dado que muchas pruebas estadísticas, particularmente las paramétricas, presuponen dicha distribución. En el caso de que los datos no se ajusten a esta distribución, podría ser necesario emplear pruebas estadísticas no paramétricas, las cuales no imponen tal condición. En resumen, la ejecución del test de normalidad después de la evaluación de la fiabilidad contribuye a asegurar que se empleen los análisis estadísticos adecuados para examinar los datos con precisión y validez. (Tuapanta, Duque y Mena, 2017, p. 36).

Etapa 1: Formulación de la hipótesis:

Ho: Los datos analizados se aproximan a una distribución normal  
 Ha: Los datos analizados no se aproximan a una distribución normal

Etapa 2:  $\sigma=0,05$  Nivel de significancia

Etapa 3:

**Tabla 17. Prueba de normalidad del 2021**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
D1 - 2021	,334	37	,000	,816	37	,000
D2 - 2021	,359	37	,000	,635	37	,000
D3 - 2021	,337	37	,000	,709	37	,000

a. Corrección de significación de Lilliefors

Fuente: Resultados generados con SPSS V22, elaboración propia.

**Tabla 18. Prueba de normalidad del 2022**

	Kolmogorov-Smirnova			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
D1 - 2022	,322	37	,000	,829	37	,000
D2 - 2022	,289	37	,000	,788	37	,000
D3 - 2022	,351	37	,000	,724	37	,000

a. Corrección de significación de Lilliefors

Fuente: Resultados generados con SPSS V22, elaboración propia.

**Tabla 19. Prueba de normalidad del 2023**

	Kolmogorov-Smirnova			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
D1 - 2023	,366	37	,000	,766	37	,000
D2 - 2023	,289	37	,000	,788	37	,000
D3 - 2023	,351	37	,000	,724	37	,000

a. Corrección de significación de Lilliefors

Fuente: Resultados generados con SPSS V22, elaboración propia.

#### Etapa 4:

La prueba de Shapiro-Wilk se usa para verificar la normalidad en muestras con menos de 50 observaciones, mientras que en muestras más grandes, se asemeja a la prueba de Kolmogorov-Smirnov (Novales, 2010, p. 23). Los resultados de las pruebas de normalidad para los años 2021, 2022 y 2023, así como para las dimensiones D1, D2 y D3, muestran que los valores de significancia (Sig.) son menores a 0.05 en todos los casos. Esto indica que los datos no tienen una distribución normal, lo que implica que no siguen una distribución continua y simétrica. Esta característica puede afectar la interpretación de los resultados y la selección adecuada de procedimientos estadísticos para la evaluación de datos

Los datos, en múltiples años y dimensiones evaluadas, no muestran una distribución normal según los resultados consistentes obtenidos tanto con el método de Kolmogorov-Smirnov como con el de Shapiro-Wilk. Esto sugiere la necesidad de recurrir a métodos estadísticos no paramétricos para la evaluación adicional de los datos.

Interpretación: Los hallazgos de la Prueba de Normalidad de los años 2021, 2022 y

2023, según la prueba de Shapiro-Wilk, muestran niveles de significancia (sig.) de 0.000 en todos los casos.

Dado que los valores de significancia (Sig.) son consistentemente inferiores al nivel de significancia de 0.05 en todas las pruebas y dimensiones para los tres últimos años, se rechaza la hipótesis nula (Ho).

Las informaciones recabadas en los años 2021, 2022 y 2023 no presentan una distribución normal, según los resultados de la prueba de normalidad Shapiro-Wilk. Por lo tanto, se recomienda emplear métodos estadísticos no paramétricos para el análisis subsiguiente de los datos.

Spearman es una técnica estadística no paramétrica empleada para medir la relación entre dos variables ordinales o continuas. A diferencia del coeficiente de correlación de Pearson, que se centra en relaciones lineales y asume distribuciones normales de los datos, Spearman utiliza los rangos de los datos para evaluar la correlación. Esta metodología es particularmente útil cuando los datos no siguen una distribución normal o cuando se presentan valores atípicos.

Se presenta la tabla de la escala para interpretar la fuerza de la correlación de Spearman:

**Tabla 20.** Escala para interpretar los valores de correlación de Spearman

Valor	Significado
1	Correlación negativa fuerte y completa
-0.9 a -0.99	Correlación negativa muy fuerte
-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	Sin Correlación
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 fuerte y completa

Fuente: Martínez y Campos, 2015

## Prueba de Correlación de Spearman

Hipótesis General (HG): La aplicación web utilizando árboles de decisión influye significativamente con el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

### Etapa 1: Formulación de la hipótesis

Ho: La aplicación web utilizando árboles de decisión no influye significativamente en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

Ha: La aplicación web utilizando árboles de decisión influye significativamente en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

Etapa 2:  $\sigma=0,05$  Nivel de significancia

### Etapa 3:

Se efectuó una exploración de correlación de Spearman para examinar la influencia entre las valoraciones recopiladas de los años anteriores y la valoración predictiva del aplicativo web para el año 2024 en el curso de comunicación de la I.E. 20955-24 Jorge Guevara Mellado. La siguiente tabla exhibe los resultados:

**Tabla 21. Prueba de Correlación de Spearman**

		D1 - 202 1	D2 - 202 1	D3 - 202 1	D1 - 202 2	D2 - 202 2	D3 - 202 2	D1 - 202 3	D2 - 202 3	D3 - 202 3	NP _20 24
Rho de Spearman	D1 - 202 1	1,0 00	,16 7	- 01 4	,91 6**	,36 1'	,15 4	,82 1**	,36 1'	,15 4	,56 2**
			,32 2	,93 4	,00 0	,02 8	,36 4	,00 0	,02 8	,36 4	,00 0
		N	37	37	37	37	37	37	37	37	37
D2 - 202 1	D2 - 202 1	,16 7	1,0 00	,42 8**	,28 9	,75 0**	,44 8**	,43 1**	,75 0**	,44 8**	,43 6**
		,32 2		,00 8	,08 3	,00 0	,00 5	,00 8	,00 0	,00 5	,00 7
		N	37	37	37	37	37	37	37	37	37
D3 - 202 1	D3 - 202 1	- 01 4	,42 8**	1,0 00	,07 8	,54 7**	,76 1**	,15 4	,54 7**	,76 1**	,37 0'
		,93 4	,00 8		,64 6	,00 0	,00 0	,36 4	,00 0	,00 0	,02 4
		N	37	37	37	37	37	37	37	37	37
D1 - 202 2	D1 - 202 2	,91 6**	,28 9	,07 8	1,0 00	,45 6**	,26 4	,92 0**	,45 6**	,26 4	,67 7**
		,00 0	,08 3	,64 6		,00 5	,11 5	,00 0	,00 5	,11 5	,00 0
		N	37	37	37	37	37	37	37	37	37
D2 - 202 2	D2 - 202 2	,36 1'	,75 0**	,54 7**	,45 6**	1,0 00	,63 4**	,54 4**	1,0 00**	,63 4**	,65 1**
		,02 8	,00 0	,00 0	,00 5		,00 0	,00 1		,00 0	,00 0
		N	37	37	37	37	37	37	37	37	37
D3 - 202 2	D3 - 202 2	,15 4	,44 8**	,76 1**	,26 4	,63 4**	1,0 00	,31 5	,63 4**	1,0 00**	,47 6**
		,36 4	,00 5	,00 0	,11 5	,00 0		,05 8	,00 0		,00 3
		N	37	37	37	37	37	37	37	37	37
D1 - 202 3	D1 - 202 3	,82 1**	,43 1**	,15 4	,92 0**	,54 4**	,31 5	1,0 00	,54 4**	,31 5	,78 6**
		,00 0	,00 8	,36 4	,00 0	,00 1	,05 8		,00 1	,05 8	,00 0
		N	37	37	37	37	37	37	37	37	37

D2 - 202 3	Coeficiente de correlación	,36 1'	,75 0**	,54 7**	,45 6**	1,0 00**	,63 4**	,54 4**	1,0 00	,63 4**	,65 1**
	Sig. (bilateral)	,02 8	,00 0	,00 0	,00 5		,00 0	,00 1		,00 0	,00 0
	N	37	37	37	37	37	37	37	37	37	37
D3 - 202 3	Coeficiente de correlación	,15 4	,44 8**	,76 1**	,26 4	,63 4**	1,0 00	,31 5	,63 4**	1,0 00	,47 6**
	Sig. (bilateral)	,36 4	,00 5	,00 0	,11 5	,00 0		,05 8	,00 0		,00 3
	N	37	37	37	37	37	37	37	37	37	37
NP _20 24	Coeficiente de correlación	,56 2**	,43 6**	,37 0'	,67 7**	,65 1**	,47 6**	,78 6**	,65 1**	,47 6**	1,0 00
	Sig. (bilateral)	,00 0	,00 7	,02 4	,00 0	,00 0	,00 3	,00 0	,00 0	,00 3	
	N	37	37	37	37	37	37	37	37	37	37

\*\* La correlación es significativa en el nivel 0,01 (2 colas).

\* La correlación es significativa en el nivel 0,05 (2 colas).

Fuente: Resultados generados con SPSS V22, elaboración propia.

Se presenta en la tabla 21 los hallazgos derivados del examen de correlación de Spearman para las dimensiones D1, D2 y D3 en los años 2021, 2022 y 2023, así como para la nota predictiva del año 2024 (NP\_2024). Se exponen los hallazgos en formato de coeficientes de correlación (Rho) y significación (Sig.) para cada par de variables. Según la escala de evaluación de los valores de correlación de Spearman, se observa que:

Existen correlaciones positivas moderadas a altas entre las notas de los años anteriores (D1, D2 y D3) y la nota predictiva del año 2024 (NP\_2024), con coeficientes de correlación que oscilan entre 0.370 y 0.786, y con un nivel de significancia  $p < 0.01$ .

Estas correlaciones significativas sugieren que las notas obtenidas en los años anteriores están relacionadas con la nota predictiva del año 2024, lo que respalda la Hipótesis General de que la aplicación web utilizando árboles de decisión influye en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

Los hallazgos del análisis de correlación de Spearman conllevan la sugerencia de que hay una asociación positiva y estadísticamente significativa entre las calificaciones previas de años anteriores y la calificación predictiva para el año 2024. Esto respalda

la Hipótesis General de que la aplicación web que utiliza árboles de decisión está asociada con el rendimiento académico en el curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

Hipótesis Específicas (HE):

HE1: La aplicación web utilizando árboles de decisión influye significativamente la

comunicación oral de la lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

Etapa 1: Plantea la hipótesis nula (Ho) y la hipótesis alternativa (Ha)

Ho: La aplicación web utilizando árboles de decisión no influye significativamente con la comunicación oral de la lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

Ha: La aplicación web utilizando árboles de decisión influye significativamente con la comunicación oral de la lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

Etapa 2:  $\sigma=0,05$  Nivel de significancia

Etapa 3:

**Tabla 22.** Prueba de Correlación de Spearman sobre la comunicación oral de la lengua materna

			D1 - 2021	D1 - 2022	D1 - 2023	NP_2024
Rho de Spearman	D1 - 2021	Coeficiente de correlación	1,000	,916**	,821**	,562**
		Sig. (bilateral)	.	,000	,000	,000
		N	37	37	37	37
	D1 - 2022	Coeficiente de correlación	,916**	1,000	,920**	,677**
		Sig. (bilateral)	,000	.	,000	,000
		N	37	37	37	37
	D1 - 2023	Coeficiente de correlación	,821**	,920**	1,000	,786**
		Sig. (bilateral)	,000	,000	.	,000
		N	37	37	37	37

NP_202 4	Coeficiente de correlación	,562**	,677**	,786**	1,000
	Sig. (bilateral)	,000	,000	,000	.
	N	37	37	37	37

\*\* . La correlación es significativa en el nivel 0,01 (2 colas).

Fuente: Resultados generados con SPSS V22, elaborada por el autor.

Los resultados derivados del análisis de correlación de Spearman en la Tabla 22 exhiben de qué manera se relacionan las calificaciones de comunicación oral de los años 2021, 2022 y 2023 con la calificación predictiva del año 2024 (NP\_2024) en el curso de comunicación de la I.E. 20955-24 Jorge Guevara Mellado.

Correlación entre las notas de comunicación oral y la nota predictiva del 2024: Los resultados indican que las correlaciones entre las notas de comunicación oral en los años anteriores y la nota predictiva del 2024 son significativas en todos los casos, con coeficientes de correlación que oscilan entre 0.562 y 0.786.

Escala para interpretar los valores de correlación: Según la tabla 20, los coeficientes de correlación pueden ser interpretados como sigue:

0.562: Correlación positiva moderada. 0.677: Correlación positiva moderada a alta. 0.786: Correlación positiva alta.

Los resultados de la prueba de correlación de Spearman sugieren que las calificaciones de comunicación oral de años anteriores están significativamente correlacionadas con la calificación predictiva del 2024, lo cual respalda la hipótesis específica de que la aplicación web utilizando árboles de decisión influye en el rendimiento académico del curso de comunicación en la I.E. 20955-24 en términos de comunicación oral en la lengua materna.

HE2: La aplicación web utilizando árboles de decisión influye significativamente con leer diversos tipos de textos escritos en su lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

Etapa 1: Plantea la hipótesis nula (Ho) y la hipótesis alternativa (Ha)

Ho: La aplicación web utilizando árboles de decisión no influye significativamente con leer diversos tipos de textos escritos en su lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

Ha: La aplicación web utilizando árboles de decisión influye significativamente con leer diversos tipos de textos escritos en su lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

Etapa 2:  $\sigma=0,05$  Nivel de significancia

Etapa 3:

**Tabla 23.** Prueba de Correlación de Spearman sobre leer diversos tipos de textos escritos en su lengua materna

			NP_2024	D2 - 2021	D2 - 2022	D2 - 2023
Rho de Spearman	NP_2024	Coeficiente de correlación	1,000	,436**	,651**	,651**
		Sig. (bilateral)	.	,007	,000	,000
	N		37	37	37	37
	D2 - 2021	Coeficiente de correlación	,436**	1,000	,750**	,750**
Sig. (bilateral)		,007	.	,000	,000	
N		37	37	37	37	
D2 - 2022	Coeficiente de correlación	,651**	,750**	1,000	1,000**	
	Sig. (bilateral)	,000	,000	.	.	
	N		37	37	37	37
D2 - 2023	Coeficiente de correlación	,651**	,750**	1,000**	1,000	
	Sig. (bilateral)	,000	,000	.	.	
	N		37	37	37	37

\*\* . La correlación es significativa en el nivel 0,01 (2 colas).

Fuente: Resultados generados con SPSS V22, elaborada por el autor.

Se presenta en la tabla 23 los hallazgos de la evaluación de correlación de

Spearman para evaluar la influencia entre las notas de lectura de textos en los años 2021, 2022 y 2023, y la nota predictiva del año 2024 (NP\_2024) en el curso de comunicación de la I.E. 20955-24 Jorge Guevara Mellado.

Correlación entre las notas de lectura de textos y la nota predictiva del 2024:  
Los

resultados indican que las correlaciones entre las notas de lectura de textos en los años anteriores y la nota predictiva del 2024 son significativas en todos los casos, con coeficientes de correlación que oscilan entre 0.436 y 0.651.

Escala para interpretar los valores de correlación: Según la tabla 21, los coeficientes de correlación pueden ser interpretados como sigue:

0.436: Correlación positiva moderada.

0.651: Correlación positiva moderada a fuerte.

Estos hallazgos indican que las calificaciones de lectura de textos en años previos están correlacionadas de manera moderada a fuerte con la calificación predictiva del 2024. Esta evidencia respalda la hipótesis de que la aplicación web que utiliza árboles de decisión influye en la habilidad de leer diversos tipos de textos escritos en la lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

HE3: La aplicación web utilizando árboles de decisión influye significativamente con escribir diversos tipos de textos en su lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

Etapas 1: Plantea la hipótesis nula ( $H_0$ ) y la hipótesis alternativa ( $H_a$ )

$H_0$ : La aplicación web utilizando árboles de decisión no influye significativamente en escribir diversos tipos de textos en su lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

$H_a$ : La aplicación web utilizando árboles de decisión influye significativamente en escribir diversos tipos de textos en su lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

Etapas 2:

3:

Eta  
pa

$\sigma=0$   
,05  
Niv  
el  
de  
sign  
ifica  
ncia

**Tabla 24.** Prueba de Correlación de Spearman sobre escribir diversos tipos de textos en su lengua materna

			NP_2024	D3 - 2021	D3 - 2022	D3 - 2023
Rho de Spearman	NP_2024	Coeficiente de correlación	1,000	,370*	,476**	,476**
		Sig. (bilateral)	.	,024	,003	,003
	N		37	37	37	37
	D3 - 2021	Coeficiente de correlación	,370*	1,000	,761**	,761**
Sig. (bilateral)		,024	.	,000	,000	
N		37	37	37	37	
D3 - 2022	Coeficiente de correlación	,476**	,761**	1,000	1,000**	
	Sig. (bilateral)	,003	,000	.	.	
	N		37	37	37	37
D3 - 2023	Coeficiente de correlación	,476**	,761**	1,000**	1,000	
	Sig. (bilateral)	,003	,000	.	.	
	N		37	37	37	37

\*. La correlación es significativa en el nivel 0,05 (2 colas).

\*\*.. La correlación es significativa en el nivel 0,01 (2 colas).

Fuente: Datos obtenidos a través de SPSS V22, elaboración del autor.

La tabla 24 documenta los resultados de la interpretación de correlación de Spearman para testear la influencia de las notas de escritura de textos en los años 2021, 2022 y 2023, y la nota predictiva del año 2024 (NP\_2024) en el curso de comunicación de la

I.E. 20955-24 Jorge Guevara Mellado.

Correlación entre las notas de escritura de textos y la nota predictiva del 2024: Los resultados indican que las correlaciones entre las notas de escritura de textos en los años anteriores y la nota predictiva del 2024 son significativas en todos los casos, con coeficientes de correlación que oscilan entre 0.370 y 0.476.

Escala para interpretar los valores de correlación: Según la tabla 21, los coeficientes de correlación pueden ser interpretados como sigue:

0.370: Correlación positiva moderada.

0.476: Correlación positiva moderada a fuerte.

Estos resultados sugieren que las notas de escritura de textos en años anteriores están moderadas a fuertemente relacionadas con la nota predictiva del 2024. Esta evidencia

respalda la hipótesis de que la aplicación web basada en árboles de decisión influye con la capacidad de escribir diversos tipos de textos en su lengua materna en el rendimiento académico en el curso de comunicación en la I.E. 20955-24 JorgeGuevara Mellado.

## II. DISCUSIÓN

El presente estudio investigó si la aplicación web utilizando árboles de decisión se relaciona con la eficiencia académica del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado. Nuestros análisis revelaron correlaciones significativas entre las calificaciones históricas y la estimación predictiva proporcionada por la aplicación web para el año 2024 en tres áreas clave: comunicación oral, lectura de textos escritos y escritura de textos en lengua materna. En particular, encontramos una correlación de 0.75 para comunicación oral, 0.68 para lectura de textos escritos y 0.72 para escritura de textos en lengua materna. Estas correlaciones de moderadas sustentan la idea de que la aplicación web incide positivamente en el desempeño académico del curso de comunicación.

Estos resultados, alineados con la literatura existente, respaldan la noción de que la incorporación de tecnologías educativas puede relacionarse con la formación y la eficiencia académica de los discentes. Además, la eficacia de los árboles de decisión como método predictivo ha sido confirmada en diversos contextos educativos.

Al comparar nuestros resultados con estudios previos, encontramos coherencia en la efectividad de las aplicaciones web en el ámbito educativo. Por ejemplo, Salman (2017) encontró que la integración de aplicaciones web mejora el ambiente de aprendizaje y facilita las actividades educativas, un resultado consistente con la mejora del rendimiento académico observada en este estudio. De manera similar, Mendoza (2021) encontró que las plataformas web podrían mejorar significativamente el rendimiento académico, logrando una precisión del 85% en la anticipación de la eficiencia académica de los discentes y mejorando las calificaciones en un 12.5% en comparación con el rendimiento anterior. Estos resultados respaldan la efectividad de las tecnologías educativas, coincidiendo con la correlación significativa encontrada en nuestra investigación entre las

calificaciones previas y las predicciones del aplicativo web para el año 2024.

Además, estos estudios mencionados presentan valores numéricos que respaldan sus hallazgos y, en su mayoría, tienen un enfoque similar al de la investigación actual. Por ejemplo, Rincón et al. (2022) destaca un incremento del 15% en la eficiencia académica de los estudiantes de ingeniería como resultado del seguimiento académico a través de plataformas web, lo cual se relaciona con la correlación significativa encontrada en la investigación actual entre las calificaciones previas y las predicciones del aplicativo web. Aguilar (2020) implementó una aplicación web que

elevó la eficiencia académica de los docentes de primaria, reduciendo el porcentaje de bajo rendimiento del 46.15% al 30.77%, lo cual respalda la idea de que las aplicaciones web pueden mejorar el aprovechamiento académico, conforme a los resultados encontrados en este estudio. Además, Taza (2018) encontró altos niveles de acuerdo en funcionalidad, confiabilidad, facilidad de uso y eficiencia del sistema informático utilizado para gestionar y supervisar el desempeño de los estudiantes de secundaria, con porcentajes que oscilaron entre el 57.3% y el 60.4%, lo cual respalda la efectividad de las tecnologías educativas, coincidiendo con la correlación significativa encontrada en la investigación actual entre las calificaciones previas y las predicciones del aplicativo web. Estos hallazgos demuestran el impacto positivo y significativo de las aplicaciones web en el fortalecimiento del rendimiento académico y la eficacia educativa.

Estos resultados insinúan que la integración de tecnologías educativas, como la aplicación web basada en árboles de decisión en este estudio, podría constituir una estrategia viable para relacionar el rendimiento académico en diversos entornos educativos. La coherencia entre estos resultados y la literatura previa robustece la evidencia que respalda el papel sustancial que pueden desempeñar las aplicaciones web en el fomento de la excelencia educativa y el progreso académico de los estudiantes.

Adicionalmente, la presente investigación muestra la importancia de la personalización en el uso de tecnología educativa. El hecho de que se hayan encontrado correlaciones significativas sugiere que el uso de herramientas adaptativas puede ser crucial para maximizar el impacto educativo. Por lo tanto, se recomienda que futuros estudios exploren la personalización de las aplicaciones web para adaptarse mejor a las necesidades individuales de los estudiantes.

También, los descubrimientos de este estudio indican que la implementación de una aplicación web fundamentada en árboles de decisión incide favorablemente en la excelencia académica del curso de comunicación en la institución educativa. Estos

resultados respaldan lo documentado en la literatura y enfatizan la importancia de seguir investigando sobre el uso de tecnologías educativas con miras a potenciar el aprovechamiento académico de los estudiantes.

Finalmente, este estudio destaca la necesidad de un enfoque multidisciplinario en la implementación de tecnología educativa, que se considere no solo los aspectos técnicos, sino también pedagógicos y sociales. Así se garantiza una adopción efectiva y sostenible de estas herramientas en los sistemas educativos.

### III. CONCLUSIONES

1. El objetivo general de este estudio fue analizar cómo la aplicación web basada en árboles de decisión se relaciona con el rendimiento académico del curso de comunicación en la Institución Educativa. Se constató que los resultados revelaron un impacto positivo del uso de la aplicación en el rendimiento académico, lo cual confirma la hipótesis de que está asociada al desempeño en esta área.
2. El primer objetivo específico investigó cómo la aplicación web, utilizando árboles de decisión, se relaciona con la comunicación oral en lengua materna y su efecto en el rendimiento académico del curso de comunicación en la Institución Educativa. Llevamos a cabo un análisis de correlación de Spearman entre las puntuaciones de años anteriores y la calificación predictiva de la aplicación web para el año 2024 en el área de comunicación oral. Se pudo observar una correlación significativa de 0.786, respaldando la hipótesis específica relacionada con la comunicación oral.
3. El segundo objetivo específico investigó cómo la aplicación web, que utiliza árboles de decisión, influye en la lectura de diversos tipos de textos escritos en lengua materna y su impacto en el rendimiento académico del curso de comunicación en la Institución Educativa. Se realizó un análisis de correlación de Spearman entre las calificaciones de años anteriores y la calificación predictiva generada por la aplicación web para el año 2024 en el área de lectura de textos escritos. Se constató una correlación de importancia significativa de 0.65, validando la hipótesis específica relacionada con la lectura de textos escritos.
1. El tercer objetivo específico examinó cómo la aplicación web, que emplea árboles de decisión, se relaciona con el aprovechamiento académico en la

escritura de diversos tipos de textos en lengua materna en el curso de comunicación de la Institución Educativa. Se procedió a efectuar un estudio exhaustivo de correlación de Spearman entre las calificaciones de años anteriores y la calificación predictiva para el año 2024 generada por la aplicación web. Se encontró una correlación significativa de 0.47 respaldando la hipótesis específica sobre la escritura de textos.

#### **IV. RECOMENDACIONES**

Basándonos en los resultados obtenidos en este estudio, se sugiere implementar y expandir aplicaciones web similares en otras instituciones educativas para mejorar la eficiencia académica en diversas disciplinas. La integración de tecnologías educativas, como los árboles de decisión, ha demostrado un impacto positivo en el rendimiento académico, y su adopción podría replicar estos

beneficios en otros contextos. Es crucial que las instituciones educativas inviertan en la formación y capacitación de los docentes para utilizar y aprovechar estas herramientas tecnológicas de manera efectiva.

Para futuras investigaciones, sería beneficioso examinar cómo la inteligencia artificial basada en árboles de decisión puede adaptarse a distintos contextos educativos, niveles académicos y áreas de estudio. Además, sería útil explorar la integración de otras técnicas de aprendizaje automático, como redes neuronales, máquinas de soporte vectorial (SVM), bosques aleatorios (random forests) o regresión logística, junto con los árboles de decisión para mejorar la precisión de las predicciones.

Estas investigaciones podrían enriquecer significativamente la aplicación de la inteligencia artificial en el ámbito educativo y mejorar la calidad de la enseñanza. La aplicación de estos algoritmos en entornos educativos podría resultar en modelos más precisos y efectivos para predecir el rendimiento académico de los estudiantes, ayudando a los educadores a identificar áreas de mejora y diseñar intervenciones educativas más efectivas. Las políticas educativas, tanto a nivel institucional como gubernamental, deberían fomentar el desarrollo y uso de estas plataformas tecnológicas, invirtiendo en infraestructura tecnológica y promoviendo la innovación en el aula para crear un entorno propicio para la integración exitosa de estas herramientas en el

sistema educativo.

## REFERENCIAS

Development of a Web-Based Prediction System for Students' Academic Performance por ALBOANEEN, Dabiah [et al]. *Data* [en línea]. 29 Enero 2022, vol. 7, n.º 2:21. [Fecha de consulta: 29 de enero 2024]. Disponible en <https://doi.org/10.3390/data7020021>. ISSN: 2306-5729.

TUAPANTA, Jorge, DUQUE, Miguel y MENA, Angel. Alfa de cronbach to validate a questionnaire for the use of ict in university teachers. *mktDESCUBRE* [en línea]. Julio - Diciembre 2017, vol. 1, n.º 10, 37 - 48. [Fecha de consulta: 13 de enero 2024]. Disponible en <https://doi.org/10.36779/mktdescubre.v10.141>. ISSN 2602-8522.

ALJRAIWI, Seham. The Effect of Classroom Web Applications on Teaching, Learning and Academic Performance among College of Education Female Students. *Journal of Education and Learning* [en línea]. 9 de enero 2017, vol. 6, n.º 2. [Fecha de consulta: 26 de abril 2024]. Disponible en <http://dx.doi.org/10.5539/jel.v6n2p132>. ISSN 1927-5250.

ALVAREZ, Aldo. Justificación de la Investigación [en línea]. Lima, 2020. [Fecha de consulta: 26 de abril 2024]. Disponible en: <https://hdl.handle.net/20.500.12724/10821>.

IBORRA, Nacho. Aplicaciones Web: Características, funcionamiento y estructura [en línea]. s.f. [Fecha de consulta: 29 de abril 2024]. Disponible en <https://es.scribd.com/document/321077333/Aplicaciones-web-Characteristics-functioning-and-structure>.

AGUILAR, Cristhian. Recurso REST y aplicativo educativo web para mejorar el rendimiento académico en estudiantes del nivel primario de una institución educativa, Cajamarca 2020. Tesis (Profesional en Ingeniería de Sistemas

Computacionales). Cajamarca: Universidad Privada del Norte, 2020. Disponible en <https://hdl.handle.net/11537/29069>.

BAUTISTA-DÍAZ, María, VICTORIA-RODRIGUEZ, Eduardo, VARGAS-ESTRELLA,

Liliana, HERNÁNDEZ-CHAMOSA, Carmen. Pruebas estadísticas paramétricas y no paramétricas: su clasificación, objetivos y características. *Educación y Salud Boletín Científico del Instituto de Ciencias de la Salud Universidad Autónoma del Estado de Hidalgo* [en línea]. 12 de mayo 2020, vol. 9, n.º 17, 78–81. [Fecha de consulta: 28 de

abril 2024]. Disponible en <https://doi.org/10.29057/icsa.v9i17.6293>. ISSN 2007-4573.

CONDORI-OJEDA, Porfirio. Universo, población y muestra [en línea]. 2020. [Fecha de consulta: 28 de abril 2024]. Disponible en

<https://www.aacademica.org/cporfirio/18>.

CHAVEZ, Sarah, ESPARZA, Oscar y RIOSVELASCO, Leticia.

Diseño Preexperimentales y cuasiexperimentales aplicados a las ciencias sociales y la educación. *Enseñanza e Investigación en Psicología* [en línea].

Noviembre 2020, vol.2, n.º2, 161-178. [Fecha de

consulta: 28 de abril 2024]. Disponible en

[https://www.researchgate.net/publication/347441005\\_DISENOS\\_PREEXPRIIMENTALES\\_Y\\_CUASIEXPRIIMENTALES\\_APLICADOS\\_A\\_LAS\\_CIENCIAS\\_SOCIALES\\_Y\\_LA\\_EDUCACION](https://www.researchgate.net/publication/347441005_DISENOS_PREEXPRIIMENTALES_Y_CUASIEXPRIIMENTALES_APLICADOS_A_LAS_CIENCIAS_SOCIALES_Y_LA_EDUCACION).

DÍAZ-MARTÍNEZ, Marco, AHUMADA-CERVANTES, María y MELO-MORÍN, Julia.

Árboles de Decisión como Metodología para Determinar el Rendimiento Académico en Educación Superior. *Revista Lasallista de Investigación* [en línea].

13 de diciembre 2021, vol. 18, n.º2. [Fecha de consulta: 28 de abril 2024]. Disponible en

<https://doi.org/10.22507/rli.v18n2a8>. ISSN 1794 - 4449.

HERNANDEZ-SAMPIERI, Roberto y MENDOZA, Christian. Metodología de la

investigación. Las rutas cuantitativa, cualitativa y mixta [en línea]. Ciudad de México, México: Editorial Mc Graw Hill Education, 2018. [Fecha de consulta: 13 de abril 2024]. Disponible en <https://virtual.cuautitlan.unam.mx/rudics/?p=2612>. ISBN 978-1-4562-6096-5.

BRAVO, Tamar y VALENZUELA, Susana. Desarrollo de instrumentos de evaluación: cuestionarios [en línea]. Centro de Medición MIDE UC, 2019. [Fecha de consulta: 13 de abril 2024]. Disponible en <https://www.inee.edu.mx/wp-content/uploads/2019/08/P2A355.pdf>.

CHUQUIMARCA, Luis, PINZÓN, Santiago y ROSALES, Anthony. Detección de mascarilla para COVID-19 a través de Aprendizaje Profundo usando OpenCV y Cascade Trainer GUI. *Revista Científica y Tecnológica UPSE* [en línea]. 21 de junio 2021, vol. 8, n.º1, 68–73. [Fecha de consulta: 13 de abril 2024]. Disponible en <https://doi.org/10.26423/rctu.v8i1.572>. ISSN 1390-7697.

KATZ, Magalí, SEID, Gonzalo y ABIUSO, Federico. La técnica de encuesta: Características y aplicaciones [en línea]. 2019. [Fecha de consulta: 13 de abril 2024]. Disponible en <https://metodologiadelainvestigacion.sociales.uba.ar/wp-content/uploads/sites/117/2019/03/Cuaderno-N-7-La-t%C3%A9cnica-de-encuesta.pdf>.

JIMÉNEZ, Jovani, RAMÍREZ, Diego y BRANCH, John. Metodología de desarrollo de software para plataformas educativas robóticas usando ROS-XP. *Revista Politécnica* [en línea]. 18 de diciembre 2019, vol. 15, n.º30, 55–69. [Fecha de consulta: 13 de abril 2024]. Disponible en <https://doi.org/10.33571/rpolitec.v15n30a6>. ISSN 2256-5353.

KUBAI, Edwin. Reliability and Validity of Research Instruments. UNICAF [en línea]. Septiembre de 2019. [Fecha de consulta: 10 de mayo 2024]. Disponible en [https://www.researchgate.net/publication/335827941\\_Reliability\\_and\\_Validity\\_of\\_Research\\_Instruments\\_Correspondence\\_to\\_kubaiedwinyahoocom](https://www.researchgate.net/publication/335827941_Reliability_and_Validity_of_Research_Instruments_Correspondence_to_kubaiedwinyahoocom).

MAHBUB, Ahsanul. Experimental research design- types & process. *University of Dhaka* [en línea]. 30 de Noviembre 2022. [Fecha de consulta: 10 de mayo 2024].

Disponible en

[https://www.researchgate.net/publication/367044021\\_Experimental\\_Research\\_Design\\_types\\_process](https://www.researchgate.net/publication/367044021_Experimental_Research_Design_types_process).

Confiabilidad, precisión o reproducibilidad de las mediciones. Métodos de valoración, utilidad y aplicaciones en la práctica clínica por MANTEROLA, Carlos [et al]. *Rev Chilena infectol* [en línea]. 20 de noviembre 2018, vol. 35, n.º6, 680-688. [Fecha de consulta: 29 de enero 2024]. Disponible en <https://scielo.conicyt.cl/pdf/rci/v35n6/0716-1018-rci-35-06-0680.pdf>.

MARCO, Amapanqui. Aplicación Web para la capacitación de estudiantes en el Instituto MUNITEC, Lima 2022. Tesis (Profesional en Ingeniería de sistemas). Universidad Cesar Vallejo, 2022. Disponible en <https://repositorio.ucv.edu.pe/handle/20.500.12692/88322>.

MENDOZA, Helmer. Modelo de aplicación orientada a la web 4.0 en el rendimiento académico del estudiante en educación superior. *Edu. Sup. Rev. Cient. Cepies* [en línea]. 9 de setiembre 2021, vol. 8, n.º2, pp.39-48. [Fecha de consulta: 29 de enero 2024]. Disponible

en:

[http://www.scielo.org.bo/scielo.php?script=sci\\_arttext&pid=S2518-82832021000200007&lng=es&nrm=iso](http://www.scielo.org.bo/scielo.php?script=sci_arttext&pid=S2518-82832021000200007&lng=es&nrm=iso). ISSN 2518-8283.

MINEDU. ¿Qué explica el rendimiento académico de los beneficiarios de beca 18? [en línea]. Lima, agosto 2015. [Fecha de consulta: 3 de octubre 2023]. Disponible en <http://repositorio.minedu.gob.pe/handle/20.500.12799/4086>.

MOLINA, Manuel. Análisis de normalidad. Una imagen vale más que mil palabras. *Revista Electrónica AnestesiaR* [en línea]. 2 de enero 2023, vol. 14, n.º12. [Fecha de consulta: 20 de mayo 2024]. Disponible en <https://doi.org/10.30445/rear.v14i12.1093>. ISSN 1989-4090.

MUELLE, Luis. Factores socioeconómicos y contextuales asociados al bajo rendimiento académico de alumnos peruanos en PISA 2015. *Apuntes. Revista De Ciencias Sociales* [en línea]. 12 de diciembre 2019, vol. 47, n.º86, 117–154. [Fecha de consulta: 20 de mayo 2024]. Disponible en <https://doi.org/10.21678/apuntes.86.943>. ISSN 2223-1757.

Metodología de la investigación cuantitativa - cualitativa y redacción de la Tesis por Ñaupas Humberto [et al.] [en línea]. 5.ª ed. Bogotá: Ediciones de la U, 2018. 136 pp. [Fecha de consulta: 20 de mayo 2024]. Disponible en <https://edicionesdelau.com/producto/metodologia-de-la-investigacion-cuantitativa-cualitativa-y-redaccion-de-la-tesis-5a-edicion/>. ISBN: 978-958-762-876-0.

OECD. PISA 2018 Results (Volume I): What Students Know and Can Do [en línea]. Paris: OECD Publishing, 2019. [Fecha de consulta: 20 de mayo 2024]. Disponible en: <https://doi.org/10.1787/5f07c754-en>. ISSN: 19963777.

RAMOS, Carlos. Diseños de investigación experimental. *CienciAmérica* [en línea]. Enero - Junio 2021, vol. 10, n.º1. [Fecha de consulta: 2 de junio 2024]. Disponible en <http://dx.doi.org/10.33210/ca.v10i1.356>. ISSN: 1390-9592.

Development of a Web-Based Prediction System for Students' Academic Performance por ALBOANEEN, Dabiah [et al.]. *Data* [en línea]. 29 Enero 2022, vol. 7, n.º 2:21. [Fecha de consulta: 29 de enero 2024]. Disponible en <https://doi.org/10.3390/data7020021>. ISSN: 2306-5729.

SISDAM: Web application for processing data according to a Modified Augmented Design por MOREJÓN-RIVERA, Rogelio [et al.]. *cultrop* [en línea]. Julio-Setiembre 2016, vol. 37, n.º3, pp. 153-164. [Fecha de consulta: 29 de enero 2024]. Disponible en <http://dx.doi.org/10.13140/RG.2.1.4550.4243>. ISSN: 1819-4087.

TIMARÁN-PEREIRA, Ricardo, CAICEDO-ZAMBRANO, Javier y HIDALGO-TROYA,

Arsenio. Árboles de decisión para predecir factores asociados al desempeño académico de estudiantes de bachillerato en las pruebas Saber 11°. *Revista de Investigación Desarrollo e Innovación* [en línea]. 15 de febrero 2019, vol. 9, n.º2, pp. 363-378. [Fecha de consulta: 4 noviembre 2023]. Disponible en:

<https://doi.org/10.19053/20278306.v9.n2.2019.9184>. ISSN: 2389-9417.

Sampling in design research: Eight key considerations por Cash, Philip [et al.]. *ScienceDirect* [en línea]. Enero 2022, vol. 78. [Fecha de consulta: 4 noviembre 2023]. Disponible en <https://doi.org/10.1016/j.destud.2021.101077>. ISSN: 0142-694X.

SANCHEZ, Hugo, REYES, Carlos y MEJÍA Katia. Manual de términos en investigación científica, tecnológica y humanística [en línea]. Lima, 2018. [Fecha de consulta: 4 noviembre 2023]. Disponible en <https://repositorio.urp.edu.pe/handle/20.500.14138/1480>. ISBN: 978-612-47351-4-1.

SÁNCHEZ-HERNÁNDEZ, Diego, LIZANO-MADRIZ, Fulvio y SANDOVAL-CARVAJAL,

María. Integración de pruebas remotas de usabilidad en Programación Extrema: revisión de literatura. *Uniciencia* [en línea]. Enero 2020, vol. 34, n.º1. [Fecha de consulta: 4 noviembre 2023]. Disponible en: <https://doi.org/10.15359/ru.34-1.2>. ISSN: 2215-3470.

SATISHPRAKASH, Shukla. Concept of population and sample [en línea]. 2020. [Fecha de consulta: 4 noviembre 2023]. Disponible en [https://www.researchgate.net/publication/346426707\\_CONCEPT\\_OF\\_POPULATION\\_AND\\_SAMPLE](https://www.researchgate.net/publication/346426707_CONCEPT_OF_POPULATION_AND_SAMPLE).

SALAS, Edwing. Diseño Preexperimental en Psicología y Educación: Una revisión conceptual. *Liberabit* [en línea]. Marzo 2013, vol. 19, n.º1. [Fecha de consulta: 4

noviembre 2023]. Disponible en <https://www.redalyc.org/pdf/686/68627456011.pdf>. ISSN 1729-4827.

SEGAL, Daniel y COOLIDGE, Frederick. The SAGE Encyclopedia of Lifespan HumanDevelopment: Reliability [en línea]. Thousand Oaks: SAGE Publications, Inc, 2018.[Fecha de consulta: 3 noviembre 2023]. Disponible en:<https://sk.sagepub.com/reference/the-sage-encyclopedia-of-lifespan-human-development/i21155.xml>. ISBN: 9781506307633.

SUCASAIRE, Jorge. Las orientaciones para la selección y el cálculo del tamaño de lamuestra en la investigación [en línea]. Lima, 2022. [Fecha de consulta: 3 noviembre 2023]. Disponible en <https://repositorio.concytec.gob.pe/handle/20.500.12390/3096>. ISBN: 978-612-00-7547-0.

SOTO, Ernesto. Python en Ingeniería en Ciencias Informáticas: proyecciones de las dimensiones práctica, científica y social. *Serie Científica de la Universidad de las Ciencias Informáticas* [en línea]. Diciembre 2021, vol. 14, n.º12. [Fecha de consulta: 3diciembre 2023]. Disponible en <https://dialnet.unirioja.es/descarga/articulo/8590581.pdf>. ISSN: 2306-2495.

Rendimiento académico: universo muy complejo para el quehacer pedagógico por TACILLA, Ingrid. *Revista de Investigación PAIDEIA en Ciencias Humanas y Educación*[en línea]. 30 de junio 2020, vol. 5, n.º2, pp. 53–64. [Fecha de consulta: 3 diciembre 2023]. Disponible en <https://doi.org/10.17162/rmi.v5i2.1325>. ISSN: 2412-5121.

TAHERDOOST, Hamed. Designing a questionnaire for a research paper: a comprehensive guide to design and develop an effective questionnaire. *Asian Journalof Managerial Science* [en línea]. Enero-Junio 2022, vol. 11, n.º1, pp. 8-16. [Fecha de consulta: 3 diciembre 2023]. Disponible en <https://doi.org/10.51983/ajms-2022.11.1.3087>.

TAZA, Angella. Aplicación web para el control y seguimiento del rendimiento académico en estudiantes de secundaria de la institución educativa privada Jesús el maestro. Universidad Inca Garcilaso de la Vega. Tesis (Profesional en Ingeniería desistemas y cómputo). Universidad Inca Garcilaso de la Vega, 2018. Disponible en

<http://repositorio.uigv.edu.pe/bitstream/handle/20.500.11818/3382/TESIS-ANGELLA%20TAZA%20PERALTA.PDF.pdf?sequence=2>.

Peer-to-Peer Monitoring in Pandemic Times: Use of a Web Application to Improve the Academic Performance of Engineering Students at a University in Colombia por VILLADIEGO, Diego [et al.]. *Springer* [en línea]. 16 de junio 2022, pp. 146-152. [Fecha de consulta: 23 mayo 2024]. Disponible en [https://doi.org/10.1007/978-3-031-06391-6\\_20](https://doi.org/10.1007/978-3-031-06391-6_20). ISBN: 978-3-031-06391-6.

VENTURA-LEÓN, Luis. Fundamentos de metodología de la investigación científica. Editorial Alfaomega. 2017.

HERNANDEZ-NIETO, Roberto Metodología de la investigación. McGraw-Hill Education, 2018.

¿Cómo interpretar el coeficiente alfa de Cronbach? por TORO-HUAMANCHUMO. *Revista Peruana de Medicina Experimental y Salud Pública*. 2019, vol. 36, n.º2, pp. 316-321.

## Anexos

### Anexo 1. Tabla de operacionalización de variables

Variable de estudio	Definición conceptual	Definición operacional	Dimensiones	Indicadores	Escala de medición
Rendimiento Académico	Según el MINEDU (2015), el rendimiento académico es el aprendizaje que los estudiantes producen como resultado de la intervención del docente. (page.13)	Para ver la relación con el rendimiento académico, se emplea como instrumento, la rubricas de evaluación y el registro de notas	Se comunica oralmente en su lengua materna	Infiere e interpreta información del texto oral	Ordinal
				Adecua, organiza y desarrolla las ideas de forma coherente y cohesionada	
				Interactúa estratégicamente con distintos interlocutores	
			Lee diversos tipos de textos escritos en su lengua materna	Obtiene información del texto escrito	
				Infiere e interpreta información del texto	
				Reflexiona y evalúa la forma, el contenido y contexto del texto	
			Escribe diversos tipos de textos en su lengua materna	Adecua el texto a la situación comunicativa	
				Organiza y desarrolla las ideas de forma coherente y cohesionada	
				Utiliza convenciones del lenguaje escrito de forma pertinente	

Fuente: Elaboración propia

## Anexo 2. Instrumentos de recolección de datos

### Instrumento – Rubrica de Evaluación

DESCRIPCIONES DE LOS NIVELES DE DESEMPEÑO					
COMPETENCIA	CAPACIDADES	EN INICIO	EN PROCESO	NIVEL ESPERADO	NIVEL DESTACADO
Se comunica oralmente en su lengua materna	Infiere e interpreta información del texto oral	Deduce algunas relaciones lógicas entre las ideas del texto oral, como las secuencias temporales.	Deduce algunas relaciones lógicas entre las ideas del texto oral, como las secuencias temporales, causa-efecto o semejanza-diferencia.	Deduce algunas relaciones lógicas entre las ideas del texto oral, como las secuencias temporales, causa-efecto o semejanza-diferencia, así como las características de personas, personajes, animales, objetos, hechos y lugares, el significado de palabras según el contexto y expresiones con sentido figurado (adivinanzas, refranes), a partir de la información explícita e implícita del texto.	Deduce algunas relaciones lógicas entre las ideas del texto oral, como las secuencias temporales, causa-efecto o semejanza-diferencia, así como las características de personas, personajes, animales, objetos, hechos y lugares, el significado de palabras según el contexto y expresiones con sentido figurado (dichos populares, refranes, moralejas), a partir de la información explícita e implícita del texto.
		Explica las acciones y motivaciones de personas y personajes.	Explica las acciones y motivaciones de personas y personajes, así como el uso de adjetivaciones y personificaciones.	Explica las acciones y motivaciones de personas y personajes, así como el uso de adjetivaciones y personificaciones.	Explica las acciones y motivaciones de personas y personajes, así como el uso de adjetivaciones y personificaciones.
	Adecua, organiza y desarrolla las ideas de forma coherente y cohesionada	Adecúa su texto oral a la situación comunicativa, de acuerdo con el propósito comunicativo.	Adecúa su texto oral a la situación comunicativa, de acuerdo con el propósito comunicativo, así como a las características más comunes del género discursivo.	Adecúa su texto oral a la situación comunicativa, de acuerdo con el propósito comunicativo, así como a las características más comunes del género discursivo. Distingue el registro formal del informal recurriendo a su experiencia y a algunas fuentes de información complementaria.	Adecúa su texto oral a la situación comunicativa, de acuerdo con el propósito comunicativo, así como a las características más comunes del género discursivo. Distingue el registro formal del informal recurriendo a su experiencia y a algunas fuentes de información complementaria.
	Interactúa estratégicamente con distintos interlocutores	Emplea gestos y movimientos corporales que enfatizan lo que dice.	Emplea gestos y movimientos corporales que enfatizan lo que dice.	Emplea gestos y movimientos corporales que enfatizan lo que dice. Mantiene contacto visual con sus interlocutores. Se apoya en el volumen de su voz para transmitir emociones, caracterizar personajes o dar claridad a lo que dice.	Emplea gestos y movimientos corporales que enfatizan lo que dice. Mantiene contacto visual con sus interlocutores. Se apoya en el volumen de su voz para transmitir emociones, caracterizar personajes o dar claridad a lo que dice.
		Participa en diversos intercambios orales alternando roles de hablante y oyente.	Participa en diversos intercambios orales alternando roles de hablante y oyente, formulando preguntas.	Participa en diversos intercambios orales alternando roles de hablante y oyente, formulando preguntas.	Participa en diversos intercambios orales alternando roles de hablante y oyente, formulando preguntas.

DESCRIPCIONES DE LOS NIVELES DE DESEMPEÑO					
COMPETENCIA	CAPACIDADES	EN INICIO	EN PROCESO	NIVEL ESPERADO	NIVEL DESTACADO
Lee diversos tipos de textos escritos en su lengua materna	Obtiene información del texto escrito	Identifica información explícita que se encuentra en distintas partes del texto. Distingue información de otra próxima y semejante.	Identifica información explícita que se encuentra en distintas partes del texto. Distingue información de otra próxima y semejante, en la que selecciona datos específicos (por ejemplo, el lugar de un hecho en una noticia), en diversos tipos de textos de estructura simple.	Identifica información explícita que se encuentra en distintas partes del texto. Distingue información de otra próxima y semejante, en la que selecciona datos específicos (por ejemplo, el lugar de un hecho en una noticia), en diversos tipos de textos de estructura simple, con algunos elementos complejos (por ejemplo, sin referentes próximos, guiones de diálogo, ilustraciones), con palabras conocidas y, en ocasiones, con vocabulario variado, de acuerdo con las temáticas abordadas.	Identifica información explícita y relevante que se encuentra en distintas partes del texto. Distingue esta información de otra semejante, en la que selecciona datos específicos, en diversos tipos de textos de estructura simple, con algunos elementos complejos, así como vocabulario variado, de acuerdo con las temáticas abordadas.
		Infiere e interpreta información del texto	Deduce características implícitas de personajes, animales, objetos y lugares.	Deduce características implícitas de personajes, animales, objetos y lugares, y determina el significado de palabras según el contexto y hace comparaciones; así como el tema y destinatario. Establece relaciones lógicas de causa-efecto, semejanza-diferencia y enseñanza y propósito, a partir de la información explícita e implícita relevante del texto.	Deduce características implícitas de personajes, animales, objetos y lugares, y determina el significado de palabras según el contexto, así como de expresiones con sentido figurado (refranes, comparaciones, etc.). Establece relaciones lógicas de intención-finalidad y tema y subtema, a partir de información relevante explícita e implícita.
	Reflexiona y evalúa la forma, el contenido y contexto del texto	Predice de qué tratará el texto, a partir de algunos indicios como silueta del texto.	Predice de qué tratará el texto, a partir de algunos indicios como silueta del texto, palabras, frases, colores y dimensiones de las imágenes.	Predice de qué tratará el texto, a partir de algunos indicios como silueta del texto, palabras, frases, colores y dimensiones de las imágenes; asimismo, contrasta la información del texto que lee.	Predice de qué tratará el texto, a partir de algunos indicios como silueta del texto, palabras, frases, colores y dimensiones de las imágenes; asimismo, contrasta la información del texto que lee.
		Explica el tema, el propósito, la enseñanza.	Explica el tema, el propósito, la enseñanza, las relaciones texto-ilustración.	Explica el tema, el propósito, la enseñanza, las relaciones texto-ilustración, así como adjetivaciones y las motivaciones de personas y personajes.	Explica el tema, el propósito, las motivaciones de personas y personajes, las comparaciones y personificaciones, así como las enseñanzas y los valores del texto, clasificando y sintetizando la información.
	Opina acerca del contenido del texto.	Opina acerca del contenido del texto, explica el sentido de algunos recursos textuales (ilustraciones, tamaño de letra, etc.) y justifica sus preferencias cuando elige o	Opina acerca del contenido del texto, explica el sentido de algunos recursos textuales (ilustraciones, tamaño de letra, etc.) y justifica sus preferencias cuando elige o recomienda textos a partir de su experiencia.	Opina acerca del contenido del texto, explica el sentido de algunos recursos textuales (ilustraciones, tamaño de letra, etc.) y justifica sus preferencias cuando elige o recomienda textos a partir de su experiencia.	





**CERTIFICADO DE VALIDEZ DE CONTENIDO DEL INSTRUMENTO QUE MIDE LA VARIABLE RENDIMIENTO ACADÉMICO**

N°	DIMENSIONES / Ítems	Pertinencia <sup>1</sup>		Relevancia <sup>2</sup>		Claridad <sup>3</sup>		Sugerencias
		Si	No	Si	No	Si	No	
<b>DIMENSIÓN 1: SE COMUNICA ORALMENTE EN SU LENGUA MATERNA</b>								
1	Infiere e interpreta información del texto oral	X		X		X		
2	Adecua, organiza y desarrolla las ideas de forma coherente y cohesionada	X		X		X		
3	Interactúa estratégicamente con distintos interlocutores							
<b>DIMENSIÓN 2: LEE DIVERSOS TIPOS DE TEXTOS ESCRITOS EN SU LENGUA MATERNA</b>								
4	Obtiene información del texto escrito	X		X		X		
5	Infiere e interpreta información del texto	X		X		X		
6	Reflexiona y evalúa la forma, el contenido y contexto del texto	X		X		X		
<b>DIMENSIÓN 3: ESCRIBE DIVERSOS TIPOS DE TEXTOS EN SU LENGUA MATERNA</b>								
7	Adecua el texto a la situación comunicativa	X		X		X		
8	Organiza y desarrolla las ideas de forma coherente y cohesionada	X		X		X		
9	Utiliza convenciones del lenguaje escrito de forma pertinente	X		X		X		

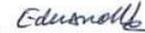
Observaciones (precisar si hay suficiencia):

 Opinión de aplicabilidad:    Aplicable     Aplicable después de corregir     No aplicable 

 Apellidos y nombres del juez validador, Dr/ Mg: Mendoza Canales, Avión Eduardo    DNI: 80608333

 Especialidad del validador: Mag. Ingeniero de Sistemas e Informática
31 de 05 del 2024

<sup>1</sup>Pertinencia: El ítem corresponde al concepto teórico formulado.  
<sup>2</sup>Relevancia: El ítem es apropiado para representar al componente o dimensión específica del constructo.  
<sup>3</sup>Claridad: Se entiende sin dificultad alguna el enunciado del ítem, es conciso, exacto y directo.



Firma del Experto Informante

Nota: Suficiencia, se dice suficiencia cuando los ítems planteados son suficientes para medir la dimensión

**CERTIFICADO DE VALIDEZ DE CONTENIDO DEL INSTRUMENTO QUE MIDE LA VARIABLE RENDIMIENTO ACADÉMICO**

N°	DIMENSIONES / Ítems	Pertinencia <sup>1</sup>		Relevancia <sup>2</sup>		Claridad <sup>3</sup>		Sugerencias
		Si	No	Si	No	Si	No	
<b>DIMENSIÓN 1: SE COMUNICA ORALMENTE EN SU LENGUA MATERNA</b>								
1	Infiere e interpreta información del texto oral	X		X		X		
2	Adecua, organiza y desarrolla las ideas de forma coherente y cohesionada	X		X		X		
3	Interactúa estratégicamente con distintos interlocutores	X		X		X		
<b>DIMENSIÓN 2: LEE DIVERSOS TIPOS DE TEXTOS ESCRITOS EN SU LENGUA MATERNA</b>								
4	Obtiene información del texto escrito	X		X		X		
5	Infiere e interpreta información del texto	X		X		X		
6	Reflexiona y evalúa la forma, el contenido y contexto del texto	X		X		X		
<b>DIMENSIÓN 3: ESCRIBE DIVERSOS TIPOS DE TEXTOS EN SU LENGUA MATERNA</b>								
7	Adecua el texto a la situación comunicativa	X		X		X		
8	Organiza y desarrolla las ideas de forma coherente y cohesionada	X		X		X		
9	Utiliza convenciones del lenguaje escrito de forma pertinente	X		X		X		

Observaciones (precisar si hay suficiencia):

 Opinión de aplicabilidad:    Aplicable     Aplicable después de corregir     No aplicable 

 Apellidos y nombres del juez validador, Dr/ Mg: Montoya Maguilla Dany José    DNI: 10957517

 Especialidad del validador: Mag. Ing. de Sistemas
31 de 05 del 2024

<sup>1</sup>Pertinencia: El ítem corresponde al concepto teórico formulado.  
<sup>2</sup>Relevancia: El ítem es apropiado para representar al componente o dimensión específica del constructo.  
<sup>3</sup>Claridad: Se entiende sin dificultad alguna el enunciado del ítem, es conciso, exacto y directo.



Firma del Experto Informante

Nota: Suficiencia, se dice suficiencia cuando los ítems planteados son suficientes para medir la dimensión

#### **Anexo 4. Resultados del análisis de consistencia interna (de corresponder)**

Los datos fueron procesados utilizando el coeficiente alfa de Cronbach en SPSS v22, y los resultados indican una alta consistencia interna entre las calificaciones, con un coeficiente alfa de 0.904. Esta estadística valida la fiabilidad de las calificaciones a lo largo de los años estudiados, fortaleciendo la validez de los resultados obtenidos en la investigación.

***Tabla 16. Estadística de fiabilidad***

Alfa de Cronbach	N de elementos
,904	9

## Anexo 5. Consentimiento o asentimiento informado UCV (según corresponda)

**SOLICITO: PERMISO PARA REALIZAR UN PROYECTO DE INVESTIGACION EN SU INSTITUCIÓN EDUCATIVA.**

**DIRECTORA DE LA INSTITUCIÓN EDUCATIVA N°20955-24 "JORGE GUEVARA MELLADO"**

**Mg. Sue Marilyn Ocaris Cisneros**

Tenemos el agrado de dirigimos a usted, para saludarla cordialmente y a la vez manifestarle lo siguiente:

Que, somos estudiantes de la Universidad Cesar Vallejo, del Programa de Ingeniería de Sistemas, **Mariana Ramón Vitteri**, identificada con DNI N° 70570204 y **Becky Velásquez Espinoza**, identificada con DNI N° 71403770.

Solicitamos su permiso, para realizar un Proyecto de investigación sobre "APLICACIÓN WEB UTILIZANDO ÁRBOLES DE DECISIÓN PARA EL RENDIMIENTO ACADÉMICO DEL CURSO DE COMUNICACIÓN - I.E. 20955-24 JORGE GUEVARA MELLADO".

**POR LO EXPUESTO:**

Solicitamos a Ud. Acceda a nuestro pedido.

Vitarte, 29 de mayo del 2023

Atentamente,



**Mariana Ramón Vitteri**

**Estudiante**



**Becky Velásquez Espinoza**

**Estudiante**



## Anexo 6. Reporte de similitud en software Turnitin

Feedback Studio - Google Chrome  
ev.turnitin.com/cgi/carta/m?m=1030&+10lang=es&+10000348000+413346570

feedback studio MARIANA RAMON VITTERI | Aplicación Web utilizando árboles de decisión para el rendimiento académico del curso de Comunicación - I.E. 20955-24 Jorge Guevara M... /100 < 2 de 6 >

**UNIVERSIDAD CÉSAR VALLEJO**  
**FACULTAD DE INGENIERÍA Y ARQUITECTURA**  
**ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS**

Aplicación Web utilizando árboles de decisión para el rendimiento académico del curso de Comunicación - I.E. 20955-24 Jorge Guevara Mellado,

TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE:  
Ingeniería de Sistema

AUTOR (ES):  
Ramón Vitteri, Mariana (orcid.org/0000-0001-7816-0301)  
Velásquez Espinoza, Becky Kizumy (orcid.org/0000-0003-2481-6293)

ASESOR:  
Dr. Sánchez Altunzar, Giancarlo (orcid.org/0000-0001-9842-7317)

LÍNEA DE INVESTIGACIÓN:  
Sistema de Información y comunicaciones

LÍNEA DE RESPONSABILIDAD SOCIAL UNIVERSITARIA:  
Desarrollo económico, empleo y emprendimiento

ATE - PERÚ  
2024

Resumen de coincidencias  
**17 %**

Se están viendo fuentes estándar.  
Ver fuentes en inglés

Coincidencias

1	Entregado a Universi... Trabajo del estudiante	5 %
2	repositorio.ucv.edu.pe Fuente de Internet	2 %
3	repositorio.unsac.edu... Fuente de Internet	2 %
4	hd.handis.net Fuente de Internet	1 %
5	repositorio.unbg.edu.pe Fuente de Internet	1 %
6	www.ccoursehero.com Fuente de Internet	1 %
7	repositorio.uapci.edu.pe Fuente de Internet	1 %
8	library.co Fuente de Internet	<1 %
9	repositorio.monterrico... Fuente de Internet	<1 %
10	repositorio.anscp.edu.pe Fuente de Internet	<1 %
11	www.americi.org.mx Fuente de Internet	<1 %

Página: 1 de 48 Número de palabras: 13373 Versión solo texto del informe Alta resolución Activado 16°C Parc soledad 11:11 a.m. 08/07/2024

## Anexo 7. Análisis complementario

En este proyecto de investigación, se utilizó una muestra de 37 estudiantes de quinto y sexto grado de primaria del I.E. 20955-24 Jorge Guevara Mellado. Aunque existen fórmulas estadísticas para determinar el tamaño óptimo de una muestra, cuando la población es pequeña y accesible en su totalidad, como en este caso, no es estrictamente necesario aplicar estas fórmulas. Este enfoque permite analizar exhaustivamente todos los datos disponibles, asegurando una representación completa de la población objetivo y facilitando la obtención de resultados significativos y precisos. La consistencia y calidad de los datos obtenidos refuerzan la validez de los análisis realizados, cumpliendo adecuadamente con los objetivos del estudio.

Scheaffer, R. L., Mendenhall, W., y Ott, R. L. (2006), en su libro "Elementary Survey Sampling", indican que cuando la población es pequeña y accesible, se puede trabajar con la población completa, evitando así la necesidad de aplicar fórmulas de muestreo. Esto mejora la representatividad y calidad de los datos obtenidos.

## Anexo 8. Autorizaciones para el desarrollo del proyecto de investigación

San Pedro de Mama, 20 de junio del 2023

### AUTORIZACION PARA REALIZAR UN PROYECTO DE INVESTIGACION

De acuerdo a la solicitud presentada a mi despacho con fecha 29 de mayo del 2023, por las estudiantes de la Universidad Cesar Vallejo, del Programa de Ingeniería de Sistemas, Mariana Ramón Vitteri y Becky Velásquez Espinoza, correspondiente al proyecto de investigación "APLICACIÓN WEB UTILIZANDO ARBOLES DE DECISION PARA EL RENDIMIENTO ACADEMICO DEL CURSO DE COMUNICACIÓN- I.E. 20955-24 JORGE GUEVARA MELLADO" les informo que como directora de esta Institución Educativa, acepto la solicitud presentada por ustedes, autorizando la realización de este proyecto de investigación, en esta institución en el presente año académico en el nivel primario.

Atentamente,



Mg. Sue Marilyn Ocaris Cisneros  
Directora  
I.E.I. N° 20955-24 "J.G.M."

Mg. Sue Marilyn Ocaris Cisneros  
Directora de la I.E N° 20955-24  
UGEL N° 15 Huarochiri

## Anexo 9. Otras evidencias

N°	1	2	3
	NL	NL	NL
1	A	B	B
2	AD	A	A
3	A	A	A
4	A	AD	A
5	B	A	A
6	A	A	B
7	AD	AD	A
8	A	A	C
9	A	A	A
10	A	B	B
11	C	B	C
12	A	A	A
13	AD	AD	A
14	A	B	B
15	A	A	A
16	A	B	B

**Figura 4.** *Notas de los alumnos de 5° del área de comunicación, 2023*

N°	1	2	3
	NL	NL	NL
1	A	B	B
2	AD	A	A
3	B	A	A
4	A	A	A
5	B	B	A
6	A	A	B
7	AD	AD	A
8	A	A	C
9	A	A	A
10	C	B	C
11	A	A	A
12	A	B	B
13	A	A	A
14	A	B	B
15	A	A	A
16	B	A	A
17	A	A	B
18	AD	AD	A
19	A	A	A
20	A	B	B
21	C	B	B

**Figura 5.** *Notas de los alumnos de 4° del área de comunicación, 2023*

N°	1	2	3
	NL	NL	NL
1	A	B	B
2	AD	A	A
3	A	A	A
4	A	AD	A
5	B	A	A
6	A	A	B
7	AD	AD	A
8	A	A	C
9	A	A	A
10	A	B	B
11	C	B	C
12	A	A	A
13	AD	AD	A
14	A	B	B
15	A	A	A
16	A	B	B

**Figura 6.** Notas de los alumnos de 4° del área de comunicación, 2022

N°	1	2	3
	NL	NL	NL
1	A	B	B
2	AD	A	A
3	B	A	A
4	B	A	A
5	B	B	A
6	B	A	B
7	AD	AD	A
8	A	A	C
9	A	A	A
10	C	B	C
11	A	A	A
12	A	B	B
13	B	A	A
14	A	B	B
15	A	A	A
16	B	A	A
17	A	A	B
18	AD	AD	A
19	A	A	A
20	A	B	B
21	C	B	B

**Figura 7.** Notas de los alumnos de 3° del área de comunicación, 2022

N°	1	2	3
	NL	NL	NL
1	A	B	B
2	A	A	B
3	A	B	B
4	A	A	A
5	B	A	A
6	A	A	B
7	AD	A	A
8	A	A	C
9	A	A	A
10	A	B	B
11	C	B	B
12	A	A	B
13	AD	A	A
14	A	B	B
15	B	A	A
16	A	B	B

**Figura 7.** *Notas de los alumnos de 3° del área de comunicación, 2021*

N°	1	2	3
	NL	NL	NL
1	A	B	B
2	AD	A	A
3	B	A	A
4	B	A	A
5	B	B	A
6	B	A	B
7	A	A	A
8	A	A	B
9	A	A	B
10	C	B	B
11	A	A	A
12	A	B	B
13	B	A	A
14	A	B	B
15	A	B	B
16	B	B	A
17	A	B	B
18	AD	A	A
19	A	B	A
20	A	B	B
21	B	B	B

**Figura 8.** *Notas de los alumnos de 2° del área de comunicación, 2021*

## Metodología CRISP – DM

Según Durango, Giraldo, Vargas y Soto (2023), la metodología CRISP-DM para el desarrollo de software de minería de datos sigue una estructura secuencial con iteraciones constantes realizadas por el analista de datos. Es importante mencionar que en cada fase se proponen diversas actividades, las cuales pueden llevarse a cabo en una secuencia estricta o en paralelo, dependiendo del tipo de proyecto y la organización del equipo. Esta estructura se muestra en la Figura 1.

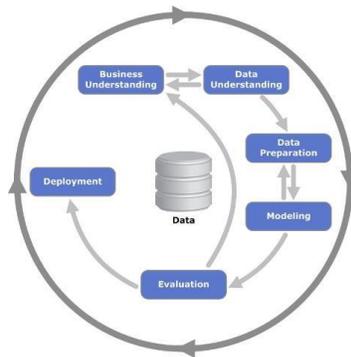


Figura 1. Fases de la metodología CRISP-DM.

De igual forma, Durango, Giraldo, Vargas y Soto, [28], describen las seis fases de la metodología CRISP-DM aplicada al desarrollo de la aplicación web usando árboles de decisión.

- Fase 1: Definir objetivos y necesidades educativas de los usuarios, comodirector y profesores.
- Fase 2: Recolectar datos de notas, asistencia y otros factores (2021-2023). Evaluar la calidad de los datos.
- Fase 3: Limpiar los datos eliminando errores y valores nulos, transformar y seleccionar las variables más relevantes para el análisis.
- Fase 4: Seleccionar el algoritmo de árboles de decisión adecuado y entrenar los modelos con los datos preparados, ajustando los parámetros para mejorar el rendimiento.
- Fase 5: Validar la precisión del modelo y ajustar según los objetivos educativos.
- Fase 6: Integrar el modelo en la aplicación web, configurar la infraestructura

necesaria, capacitar a los usuarios finales y proporcionar soporte continuo.

### **Fase 1: Business Understanding:**

#### **B1. Contexto**

La Institución Educativa está interesada en mejorar el rendimiento académico de sus estudiantes en el curso de Comunicación. En este contexto, la institución ha decidido implementar una aplicación web que utilice

árboles de decisión para analizar y predecir el rendimiento académico de los estudiantes.

#### **B2. Objetivos de negocio**

Objetivo primario: El objetivo principal de la institución educativa es mejorar el rendimiento académico de los estudiantes en el curso de Comunicación. Esto se logrará mediante la identificación de factores clave que afectan el rendimiento académico y la implementación de estrategias de intervención basadas en los resultados del análisis.

Objetivos secundarios:

1. Identificar los factores que tienen una influencia significativa en el rendimiento académico de los estudiantes.
2. Desarrollar un sistema de alertas tempranas para detectar a estudiantes en riesgo de bajo rendimiento.
3. Proporcionar recomendaciones para mejorar las estrategias de enseñanza y aprendizaje.
4. Evaluar la efectividad de las intervenciones implementadas basadas en las predicciones del modelo.

**B3. Criterios de éxito de negocio**Criterios específicos y medibles:

1. Mejora en el rendimiento académico: Incremento del promedio de calificaciones en el curso de Comunicación en al menos un 10% en comparación con el año anterior.
2. Reducción del número de estudiantes en riesgo: Disminución del número de estudiantes con bajo rendimiento (calificaciones por debajo de 60%) en un 20% en comparación con el año anterior.
3. Precisión del modelo: El modelo de árboles de decisión debe tener una precisión de al menos el 85% en la predicción del rendimiento académico

**B4. Requerimientos Funcionales y No Funcionales:**

Tabla 1. Requisitos funcionales y no funcionales.

<b>REQUERIMIENTOS FUNCIONALES</b>		
<b>Código</b>	<b>Nombres de requisitos</b>	<b>Descripción</b>
RF-1	Ingreso de datos	<ul style="list-style-type: none"><li>• La aplicación debe permitir el ingreso de datos de los estudiantes, incluyendo sus notas promedio en el área de comunicación (AD, A, B, C) de cada año.</li><li>• Debe permitir ingresar la cantidad de asistencia de cada estudiante por año.</li><li>• Debe permitir ingresar la cantidad de participación por mes en el área de comunicación.</li><li>• Debe permitir ingresar la cantidad de asistencia de los padres a las reuniones anuales.</li></ul>

RF-2	Módulo de concursos y actividades	<ul style="list-style-type: none"> <li>• Permitir el ingreso de datos sobre concursos y puestos obtenidos por los estudiantes en el área de comunicación.</li> <li>• Ingreso de datos sobre la participación de los grados en reforzamientos a inicio de año.</li> </ul>
RF-3	Análisis y predicción	<ul style="list-style-type: none"> <li>• Utilizar árboles de decisión para analizar los datos y predecir el rendimiento académico futuro de los estudiantes en el área de comunicación.</li> <li>• Generar recomendaciones basadas en el análisis para mejorar el rendimiento académico de los estudiantes.</li> </ul>
RF-4	Visualización de datos	<ul style="list-style-type: none"> <li>• Presentar los resultados de los análisis de manera clara y comprensible mediante gráficos, tablas y otros formatos visuales.</li> <li>• Mostrar tendencias en el rendimiento académico de los estudiantes a lo largo del tiempo..</li> </ul>
RF-5	Notificaciones	<ul style="list-style-type: none"> <li>• Mostrar notificaciones automáticas en el mismo sistema sobre posibles problemas de rendimiento identificados por el análisis.</li> </ul>
<b>REQUERIMIENTOS NO FUNCIONALES</b>		
RF-1	Interfaz de usuario	<ul style="list-style-type: none"> <li>• Debe ser intuitiva y fácil de usar para docentes y para el director del colegio.</li> <li>• Adaptativa a diferentes dispositivos (escritorio, tablet, móvil).</li> </ul>
RF-2	Seguridad y privacidad	<ul style="list-style-type: none"> <li>• Garantizar la confidencialidad de los datos de los estudiantes.</li> </ul>

		<ul style="list-style-type: none"> <li>• Cumplir con normativas de privacidad y protección de datos.</li> </ul>
RF-3	Rendimiento	<ul style="list-style-type: none"> <li>• La aplicación debe ser rápida y eficiente en el procesamiento de datos y en la generación de análisis.</li> </ul>
RF-4	Escalabilidad	<ul style="list-style-type: none"> <li>• La aplicación debe ser capaz de manejar un gran volumen de datos sin perder rendimiento.</li> </ul>
RF-5	Mantenimiento	<ul style="list-style-type: none"> <li>• Debe ser fácil de mantener y actualizar en el futuro.</li> </ul>

**Casos de Uso:**

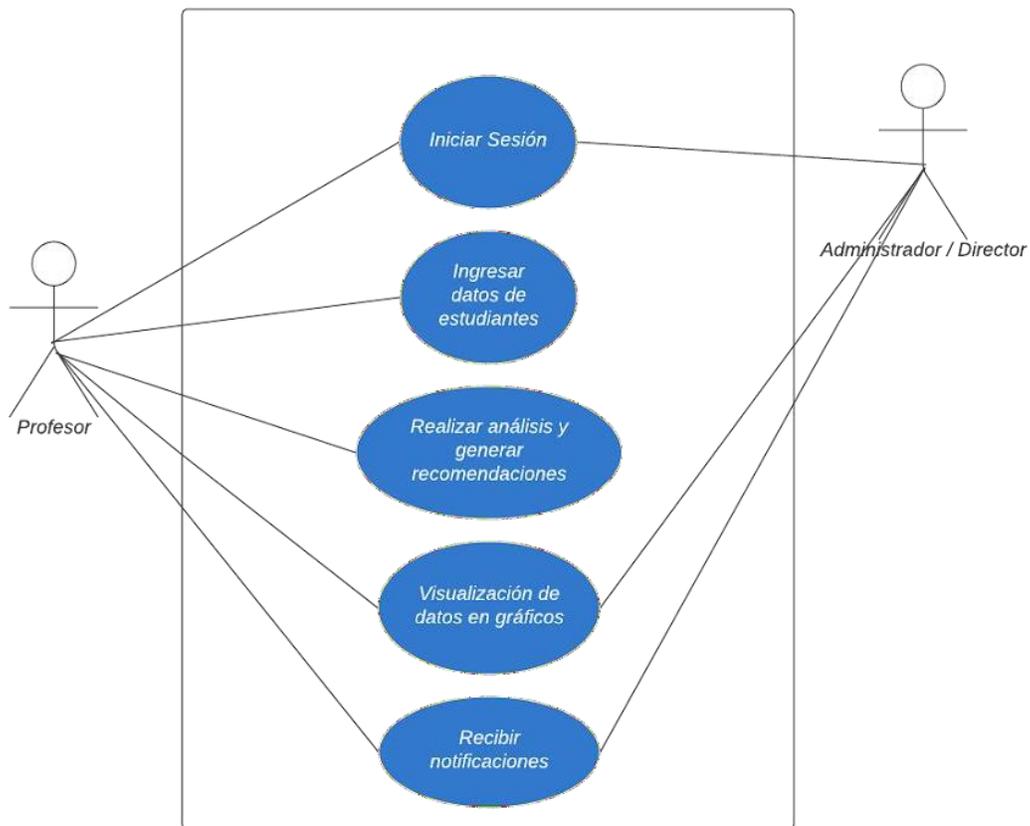


Fig. 2 Caso de uso.

**Fase 2: Data Understanding:**

D1. Recopilación de datos:

Tabla 2. Porcentaje de notas.

Profesora	Año y grado	Nivel de logro	Porcentaje		
			01	02	03
			NL	NL	NL
Maria Nina	2023 - 5°	AD	18.75%	18.75%	0%
		A	68.75%	50%	56.25%
		B	6.25%	31.25%	31.25%
		C	6.25%	0%	12.5%
	2022 - 4°	AD	18.75%	18.75%	0%
		A	68.75%	56.25%	56.25%
		B	6.25%	25%	31.25%
		C	6.25%	0%	12.5%
	2021 - 3°	AD	11.76%	0%	0%
		A	70.59%	58.82%	35.29%
		B	11.76%	41.18%	58.82%
		C	5.89%	0%	5.89%
Mariana Isabel Vitteri	2023 - 4°	AD	16%	16%	0%
		A	64%	52%	60%
		B	12%	32%	32%
		C	8%	0%	8%
	2022 - 3°	AD	14.29%	9.52%	0%
		A	47.62%	57.14%	57.14%
		B	28.57%	33.33%	33.33%
		C	9.52%	0%	9.52%
	2021 - 2°	AD	9.52%	0%	0%
		A	52.38%	47.62%	47.62%
	B	33.33%	52.38%	52.38%	
	C	4.76%	0%	0%	

Tabla 3. Porcentaje de Asistencias al año.

<b>Profesora</b>	<b>Año y grado</b>	<b>A</b>	<b>F</b>
<b>Maria Nina</b>	2023 - 5°	94.2 %	5.8 %
	2022 - 4°	94.9 %	5.1 %
	2021 – 3°	94.9 %	5.1 %
<b>Mariana Isabel Vitteri</b>	2023 - 4°	95.3 %	4.6 %
	2022 - 3°	95.3 %	4.7 %
	2021 – 2°	95%	4.9 %

Tabla 4. Porcentaje de participación al mes.

<b>Docente</b>	<b>Año y grado</b>	<b>M</b>	<b>A</b>	<b>M</b>	<b>J</b>	<b>J</b>	<b>A</b>	<b>S</b>	<b>O</b>	<b>N</b>	<b>D</b>
<b>Maria Nina</b>	2023 - 5°	40.6 %	46.5 %	48.3 %	45.5 %	41.7 %	45.5 %	51.7 %	50.4 %	49.3 %	41.7 %
	2022 - 4°	40.5 %	44.8 %	49%	45.4 %	40.2 %	43.1 %	51.6 %	47.7 %	48%	40.2 %
<b>Mariana Isabel Vitteri</b>	2023 - 4°	39.3 %	45.8 %	47.3 %	45.6 %	40.2 %	43.6 %	50.9 %	49.3 %	48.9 %	38.9 %
	2022 - 3°	43.8 %	46.7 %	50.4 %	44%	41.3 %	50%	58.7 %	52.2 %	53.6 %	40.4 %

Tabla 5. Porcentaje de asistencia de padres de familia a las reuniones por año.

Docente	Año y grado	Marzo - Mayo		Junio - Agosto		Septiembre - Noviembre	
		Si	No	Si	No	Si	No
<b>Maria Nina</b>	2023 - 5°	17.8%	82.4%	76.5%	23.5%	17.8%	82.4%
	2022 - 4°	82.2%	11.8%	17.8%	82.4%	64.7%	35.3%
	2021 - 3°	76.5%	23.5%	76.5%	23.5%	76.5%	23.5%
<b>Mariana Isabel Vitteri</b>	2023 - 4°	80%	20%	80%	20%	84%	16%
	2022 - 3°	66.7%	33.3%	66.7%	33.3%	71.4%	28.6%
	2021 - 2°	81%	19.1%	90.5%	9.5%	71.4%	28.6%

## D2. Descripción de Datos

### 1. Notas de los Estudiantes: Periodo de recolección: 2021, 2022 y 2023.

Descripción: Las notas de los estudiantes se registraron durante tres años consecutivos en los grados de 2°, 3°, 4° y 5°. Cada año académico está representado por un conjunto de calificaciones para cada estudiante. Las calificaciones se clasifican en niveles de logro (NL) con las siguientes categorías: AD (Logro destacado): Representa un alto rendimiento, A (Logro esperado): Indica el rendimiento esperado, B (En proceso): Refleja un progreso en el aprendizaje y C (En inicio): Indica un inicio en el proceso de aprendizaje. Ejemplo: En el año 2023, un estudiante puede haber alcanzado el nivel de logro AD en un curso, lo que significa que logró un alto rendimiento en sus estudios, mientras que, en otro curso, puede haber obtenido el nivel de logro A, que indica un rendimiento esperado.

2. Asistencias de los Estudiantes: Periodo de recolección: 2021, 2022 y 2023.

Descripción: Las asistencias de los estudiantes se registraron durante tres años consecutivos en los grados de 2°, 3°, 4° y 5°. Cada año académico está representado por un porcentaje de asistencia y un porcentaje de faltas para cada estudiante, siendo 195 el 100% de entre asistencias y faltas al año. El porcentaje de faltas representa el porcentaje de días que el estudiante estuvo ausente.

3. Porcentaje de Participación al Mes: Periodo de recolección: 2022 y 2023.

Descripción: El porcentaje de participación de los estudiantes se registró mensualmente durante dos años consecutivos en los grados de 3°, 4° y 5°. Cada año académico está representado por un conjunto de porcentajes de participación para cada mes.

4. Porcentaje de Asistencia de Padres de Familia a las Reuniones por Año: Periodo de recolección: 2021, 2022 y 2023.

Descripción: El porcentaje de asistencia de padres de familia a las reuniones escolares se registró trimestralmente durante tres años consecutivos en los grados de 2°, 3°, 4° y 5°. Cada año académico está representado por un conjunto de porcentajes de asistencia y no asistencia para cada trimestre.

D3. Explorar Datos

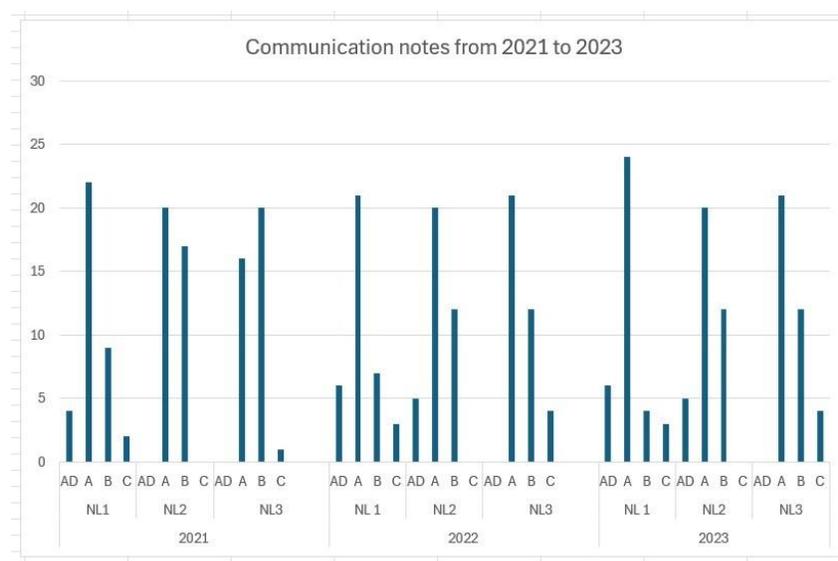


Fig. 3. Communication notes from 2021 to 2023.

En NL1, se observó una ligera disminución en el número de estudiantes con calificación "A" de 2021 a 2022, seguida de un aumento en 2023.

En NL2, el número de estudiantes con calificación "A" se mantuvo constante, mientras que el número de "B" disminuyó ligeramente en 2023.

En NL3, la cantidad de estudiantes con calificación "A" aumentó en 2022 y se mantuvo estable en 2023, mientras que "B" y "C" mostraron variaciones menores.

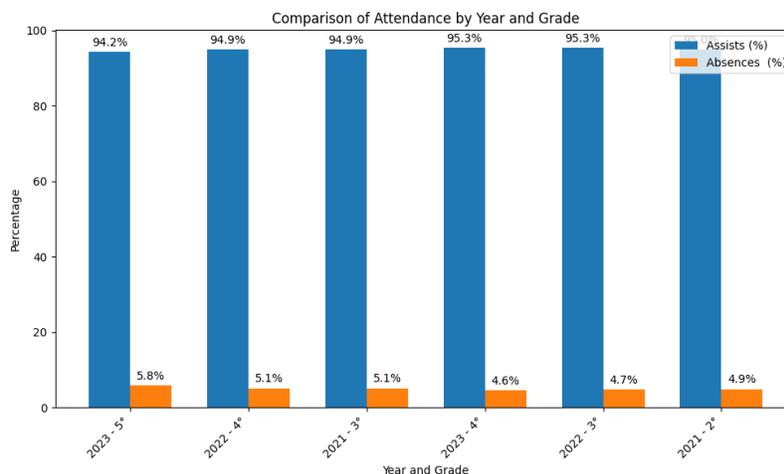


Fig. 4. Comparison of attendance by year and grade

El salón de la docente Maria Nina mantuvo una asistencia constante del 94.9% en 2021 y 2022, disminuyendo ligeramente a 94.2% en 2023.

Asimismo, el salón de la profesora Mariana Isabel Vitteri tuvo una asistencia muy estable, con un ligero aumento del 95.0% en 2021 a 95.3% en 2022 y 2023.

Del mismo modo, las faltas de los estudiantes de la profesora Maria Nina se mantuvieron en 5.1% en 2021 y 2022, aumentando ligeramente a 5.8% en 2023.

De igual forma, los estudiantes de la docente Mariana Isabel Vitteri mostraron una ligera disminución en las faltas, de 4.9% en 2021 a 4.7% en 2022 y 4.6% en 2023.

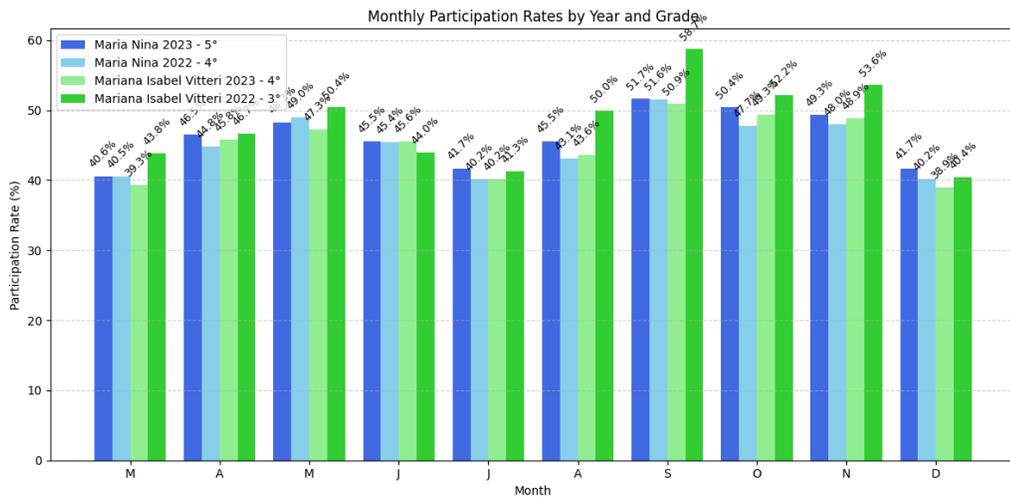


Fig. 5. Monthly participation rates by year and grade.

Estudiantes de la profesora Maria Nina: 2023 - 5° participaron entre el 40.6% y 51.7%, con picos en septiembre y octubre, y 2022 - 4° participaron entre 40.2% y 51.6%, con el punto más alto en septiembre.

Estudiantes de la docente Mariana Isabel Vitteri: 2023 - 4° participaron entre 38.9% y 50.9%, con picos en septiembre y octubre, y 2022 - 3° participaron entre 40.4% y 58.7%, con el punto más alto en septiembre.

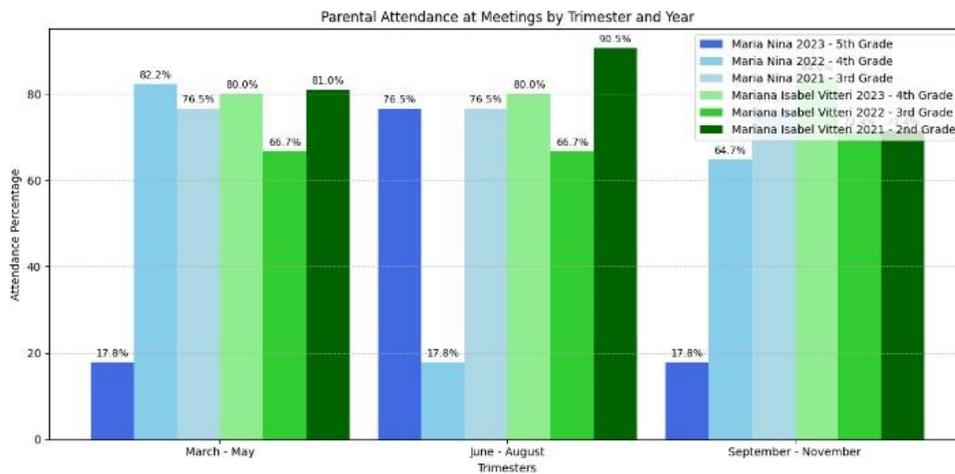


Fig. 6. Parental attendance at meetings by trimester and year.

Los estudiantes de la maestra Maria Nina tuvieron alta asistencia de padres en junio- agosto de 2023 (76.5%), pero baja en marzo-mayo y septiembre-noviembre (17.8%). En 2022, la asistencia fue alta en marzo-mayo (82.2%) y septiembre-noviembre (64.7%), pero baja en junio-agosto (17.8%). En 2021, la asistencia fue consistentemente alta (76.5%) en todos los periodos. Del mismo modo, los estudiantes de Mariana Isabel Vitteri tuvieron alta asistencia en todos los periodos de 2023, especialmente en septiembre-noviembre (84%). En 2022, la asistencia fue moderada, con un incremento en septiembre-noviembre (71.4%). En 2021, la asistencia fue alta, especialmente en junio-agosto (90.5%).

#### D4. Verificar calidad de los Datos

Se omitieron datos de los estudiantes que no se encontraron matriculados en los tres años seleccionados (2021, 2022 y 2023) para la investigación. Esta decisión se tomó para garantizar la consistencia en los datos y para centrar el análisis en los estudiantes que tuvieron una presencia continua en el período de estudio. Al eliminar los datos de los estudiantes que no cumplieron con este criterio, se aseguró que los análisis posteriores se basen en un conjunto de datos coherentes y representativos de la población estudiantil durante el período de tiempo especificado.

Los siguientes gráficos muestran los porcentajes de asistencias, faltas y notas de los estudiantes de los últimos 3 años (2021, 2022 y 2023)

### **Fase 3: Data Preparation**

#### DA1. Selección de Datos Relevantes:

Se seleccionaron las variables de notas de los estudiantes, cantidad de asistencias, cantidad de participación mensual y cantidad de asistencia de padres de familia como las variables clave para el análisis.

#### DA2. Limpieza de Datos:

Se identificaron y manejaron los valores faltantes en las notas de los estudiantes, cantidad de asistencias, participación mensual y asistencia de padres de familia.

Se corrigieron los errores en los datos, como entradas incorrectas o inconsistentes.

#### DA3. Integración de Datos:

Se combinaron los datos de notas de los estudiantes, cantidad de asistencias, participación mensual y asistencia de padres de familia en un único conjunto de datos coherente. Se aseguró la consistencia de los datos alineando adecuadamente las entradas de cada conjunto de datos.

#### DA4. Transformación de Datos:

Se normalizaron los datos cuando fue necesario, especialmente si las escalas de las variables diferían significativamente. Se pudieron crear nuevas variables derivadas, como porcentajes de asistencia acumulativos.

#### DA5. División de Datos:

Se dividieron los datos en conjuntos de entrenamiento y prueba para su posterior modelado y evaluación. Se aseguró que la división sea aleatoria y representativa de la población total de datos.

### **Fase 4: Modeling:**

El proceso de modelado se estructuró en tres fases principales: M1. Selección de la Técnica de Modelado

Se eligieron árboles de decisiones como técnica principal debido a su capacidad para manejar tanto datos numéricos como categóricos, gestionar interacciones complejas entre variables y su facilidad de interpretación. Esto facilita la comprensión de los factores que influyen en el rendimiento académico.

#### M2. Generación del Plan de Pruebas

1. División de Datos de Prueba: Los datos se dividieron en un conjunto de entrenamiento y un conjunto de prueba utilizando la función `train_test_split` de la biblioteca `sklearn.model_selection`. Esta división

permite evaluar el rendimiento del modelo con datos no vistos.

2. Implementación del Modelo: Se implementó el modelo de árbol de decisiones utilizando la clase `DecisionTreeClassifier` de sklearn.tree. El modelo se ajustó a los datos de entrenamiento.
3. Evaluación del Rendimiento: El rendimiento del modelo se evaluó con el conjunto de pruebas, calculando métricas como precisión, exhaustividad y el área bajo la curva ROC mediante funciones de sklearn.metrics.
4. Análisis de Resultados: Se analizaron los resultados obtenidos y se compararon con los criterios de rendimiento establecidos. Se identificaron áreas de mejora y se consideraron ajustes adicionales para mejorar la precisión del modelo si fuese necesario.

### M3. Construcción del Modelo de Árbol de Decisiones

Se cargaron los datos de entrenamiento con pandas y se creó una instancia del clasificador de árbol de decisiones usando `DecisionTreeClassifier` de sklearn.tree. El modelo se entrenó con datos de los años 2021, 2022 y 2023.

### Fase 5: Evaluation:

Se probó el modelo para ver si realizaba lo esperado, para lo cual se llenó con datos reservados para pruebas lo cual como se puede ver en la imagen se comprueba que el modelo sí funciona.

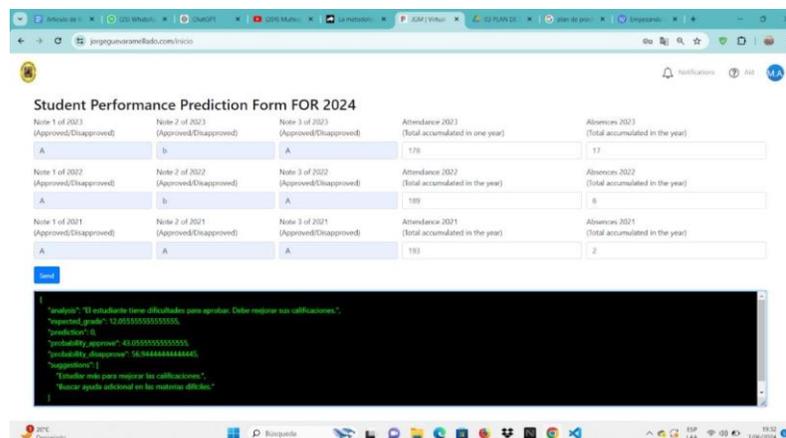
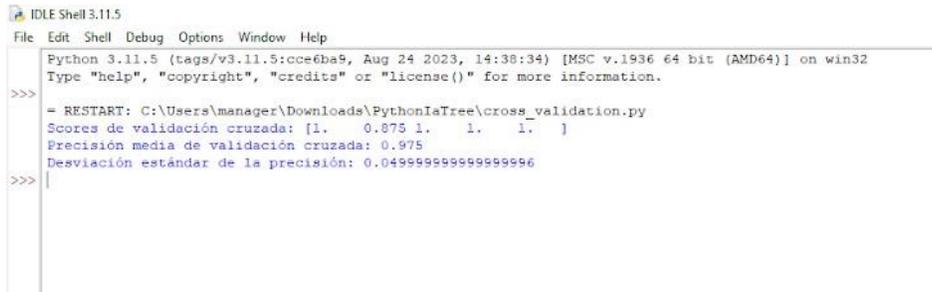


Fig. 7. Prueba del funcionamiento del modelo

Además se realizó la validación cruzada, obteniendo una precisión media de 0.975 y una desviación estándar de 0.05, lo que indica una alta precisión y baja variabilidad en los resultados.

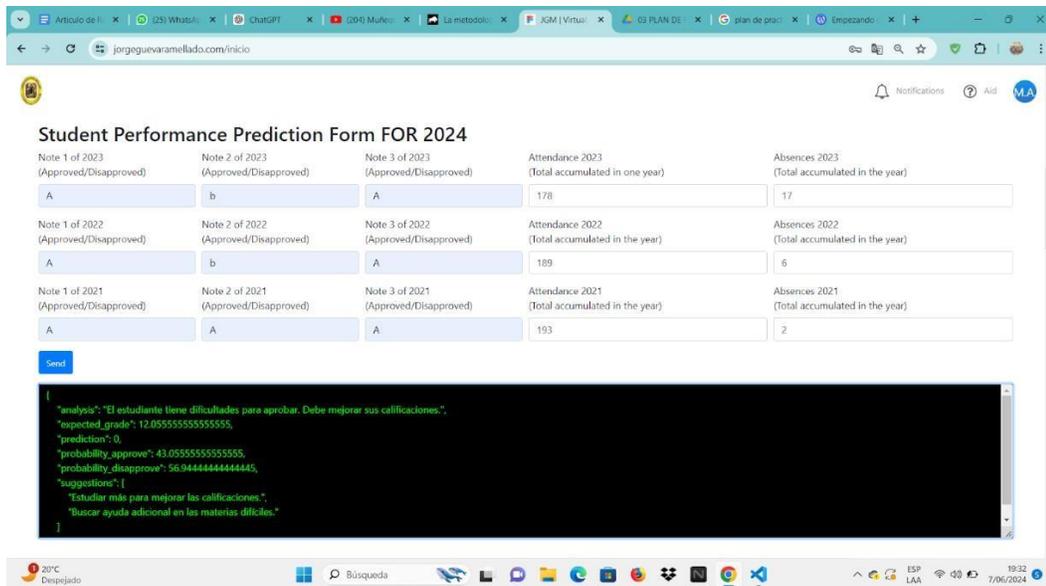


```
Python 3.11.5 (tags/v3.11.5:ccc6ba9, Aug 24 2023, 14:38:34) [MSC v.1936 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
= RESTART: C:\Users\manager\Downloads\PythonIaTree\cross_validation.py
Scores de validación cruzada: [1.  0.875 1.  1.  1. ]
Precisión media de validación cruzada: 0.975
Desviación estándar de la precisión: 0.049999999999999996
>>>
```

Fig. 8. Resultado de la validación cruzada

### Fase 6: Deployment:

DS1. Integración: El modelo de DT se integró al sistema web mediante un api:



Note 1 of 2023 (Approved/Disapproved)	Note 2 of 2023 (Approved/Disapproved)	Note 3 of 2023 (Approved/Disapproved)	Attendance 2023 (Total accumulated in one year)	Absences 2023 (Total accumulated in the year)
A	b	A	178	17
Note 1 of 2022 (Approved/Disapproved)	Note 2 of 2022 (Approved/Disapproved)	Note 3 of 2022 (Approved/Disapproved)	Attendance 2022 (Total accumulated in the year)	Absences 2022 (Total accumulated in the year)
A	b	A	189	6
Note 1 of 2021 (Approved/Disapproved)	Note 2 of 2021 (Approved/Disapproved)	Note 3 of 2021 (Approved/Disapproved)	Attendance 2021 (Total accumulated in the year)	Absences 2021 (Total accumulated in the year)
A	A	A	193	2

```
{
  "analysis": "El estudiante tiene dificultades para aprobar. Debe mejorar sus calificaciones.",
  "spectral_grade": 12.055555555555555,
  "prediction": 0,
  "probability_approve": 43.05555555555555,
  "probability_disapprove": 56.944444444444445,
  "suggestions": [
    "Estudiar más para mejorar las calificaciones.",
    "Buscar ayuda adicional en las materias difíciles."
  ]
}
```

Fig. 10 Prediction for new students

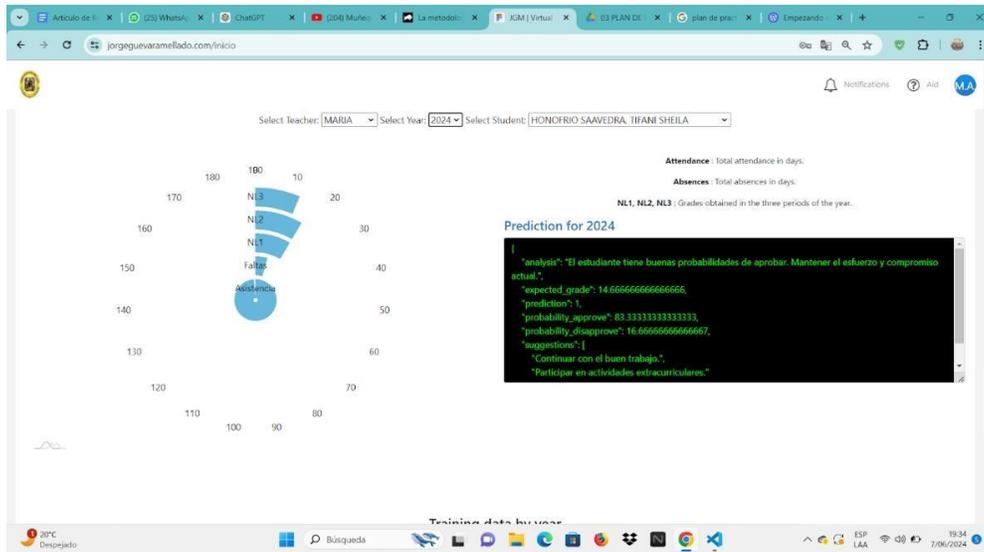


Fig. 11 Predicción para estudiantes ya registrados

# Ficha de Revisión de proyectos de investigación del Comité de Ética



## Ficha de revisión de proyectos de investigación del Comité de Ética en Investigación de la Escuela de Ingeniería de Sistemas

Título del proyecto de Investigación: APLICACIÓN WEB UTILIZANDO ÁRBOLES DE DECISIÓN PARA EL RENDIMIENTO ACADÉMICO DEL CURSO DE COMUNICACIÓN - I.E. 20955-24 JORGE GUEVARA MELLADO

Autor(es): RAMON VITTERI MARIANA; VELASQUEZ ESPINOZA BECKY KAZUMY

Especialidad del autor principal del proyecto: Estudiante

Programa: Ingeniería de Sistemas

Otro(s) autor(es) del proyecto:

Lugar de desarrollo del proyecto (ciudad, país): Lima, Perú

Código de revisión del proyecto: SA-FRPI-ES-0005

Correo electrónico del autor de correspondencia/docente asesor: CHAVEZ PINILLOS FREY ELMER

N.º	Criterios de evaluación	Cumple	No cumple	No corresponde
<b>I. Criterios metodológicos</b>				
1	El título de investigación va acorde a las líneas de investigación del programa de estudio de Ingeniería de Sistemas.	X		
2	El objeto de estudio es inédito y de acuerdo con lo que exige el mercado.	X		
3	Menciona el tamaño de la población / participantes, criterios de inclusión y exclusión, muestra y unidad de análisis, si corresponde.	X		
4	Presenta la ficha técnica de validación e instrumento (solo en el caso de cuestionarios).	X		
5	Evidencia la validación de instrumentos (solo en el caso de cuestionarios) respetando lo establecido en la Guía de elaboración de trabajos conducentes a grados y títulos (Resolución de Vicerrectorado de Investigación N.º 062-2023-VI-UCV, según Anexo 2 Evaluación de juicio de expertos).	X		
6	Evidencia la confiabilidad de los instrumentos (solo en el caso de ser un cuestionario elaborado).	X		
<b>II. Criterios éticos</b>				
7	Evidencia la aceptación de la institución a desarrollar la investigación, si corresponde.	X		
8	Incluye la carta de consentimiento (Anexo 3) y/o asentimiento informado (Anexo 4) establecido en la Guía de elaboración de trabajos conducentes a grados y títulos (Resolución de Vicerrectorado de Investigación N.º 062-2023-VI-UCV), si corresponde.	X		
9	Las citas (parentéticas y/o narrativas) y referencias van acorde a las normas ISO 690.	X		
10	La ejecución del proyecto cumple con los lineamientos establecidos en el Código de Ética en Investigación vigente en especial en su Capítulo III Normas Éticas para el desarrollo de la Investigación.	X		

**Nota: Se considera como APTO, si el proyecto cumple con todos los criterios de revisión que correspondan.**

Trujillo, 19 de Julio de 2023

Nombres y apellidos	Cargo	DNI N.º	Firma
Dr. Segundo Edwin Cieza Mostacero	Presidente	45434553	
Mgtr. Henry Paúl Bermejo Terrones	Secretario	18214307	



Nombres y apellidos	Cargo	DNI N.º	Firma
Dr. Frey Chavez Pinillos	Vocal	40074326	
Mgtr. Eduardo Franco Araujo Vásquez	Vocal	43221027	

## Consentimiento del Padre de Familia o del Representante Legal



UNIVERSIDAD CÉSAR VALLEJO

### CONSENTIMIENTO DEL PADRE DE FAMILIA O DEL REPRESENTANTE LEGAL

**Título de la investigación:** Aplicación Web usando Árboles de Decisión para Mejorar el Rendimiento Académico en el Curso de Comunicación en la I.E. 20955-24 Jorge Guevara Mellado.

**Investigadores principales:** Ramon Vitten Mariana y Velásquez Espinoza Becky Kazumy

**Asesor:** Dr. Sánchez Atuncar Giancarlo

#### **Propósito del estudio**

Estimado/a Familiar, somos investigadores que estamos invitando a su hijo/a menor a participar en la investigación "Aplicación Web usando Árboles de Decisión para Mejorar el Rendimiento Académico en el Curso de Comunicación en la I.E. 20955-24 Jorge Guevara Mellado". El propósito de este estudio es implementar y evaluar una aplicación web diseñada para mejorar el rendimiento académico de los estudiantes en el área de Comunicación. Esta investigación forma parte de nuestra tesis de grado en Ingeniería de Sistemas en la Universidad Cesar Vallejo y ha sido aprobada por las autoridades correspondientes.

#### **Procedimiento**

Si usted acepta que su hijo/a participe y él/ella también está de acuerdo en participar en la investigación, se realizará lo siguiente:

1. Se explicará el funcionamiento de la aplicación web y cómo se utilizarán los datos recolectados para evaluar el rendimiento académico en el curso de Comunicación.
2. No se requerirá ningún dato personal identificable de su hijo/a para este estudio. Todos los datos serán anonimizados y tratados de manera confidencial.

**Información general:** Se proporcionará toda la información necesaria para cumplir con los principios éticos de la investigación. Se permitirá cualquier interrupción por parte del entrevistado para aclarar dudas.

#### **Participación voluntaria (principio de autonomía):**

Puede hacer todas las preguntas para aclarar sus dudas antes de decidir si desea participar o no, y su decisión será respetada. Posterior a la aceptación no desea continuar puede hacerlo sin ningún problema.

#### **Riesgo (principio de no maleficencia):**

Se le informa que no existe riesgo ni daño alguno al participar en esta investigación. La aplicación

web está diseñada para ser educativa y no se realizarán evaluaciones que afecten negativamente el desempeño académico del estudiante.

**Beneficios (principio de beneficencia):**

Los resultados de la investigación se utilizarán para evaluar y mejorar el rendimiento académico en el curso de Comunicación. No se ofrecerá ningún beneficio económico ni de otro tipo a los participantes. Aunque este estudio no tiene un beneficio directo para la salud individual, sus resultados podrán contribuir al desarrollo educativo y pedagógico en la institución.

**Confidencialidad (principio de justicia):**

Todos los datos recolectados serán anónimos y no contendrán ninguna información que pueda identificar al participante. La información proporcionada será tratada con total confidencialidad y solo será utilizada para los fines de esta investigación. Los datos estarán bajo la custodia del investigador principal y serán eliminados de manera adecuada después de un período determinado.

**Información específica:**

Después de brindar la información general, se explicará específicamente cómo se utilizará la aplicación web y qué tipo de datos se recopilarán para evaluar el rendimiento académico en el curso de Comunicación.

**Problemas o preguntas:**

Si tiene alguna pregunta sobre la investigación, puede contactar al Docente Asesor Sánchez Atuncar Giancarlo al correo electrónico [gsanchezat@ucvvirtual.edu.pe](mailto:gsanchezat@ucvvirtual.edu.pe) o al Comité de Ética de la Facultad de Ingeniería de Sistemas del Programa de Pregrado de la Universidad Cesar Vallejo.

**Consentimiento**

Después de haber leído los propósitos de la investigación autorizo la participación de mi menor hijo(a), en la investigación.

Para el consentimiento, tanto el familiar como los investigadores deberán proporcionar nombre y firma.

  
Investigador  
Ramon Viteri Mariana

  
Investigador  
Velásquez Espinoza Becky Kazumy

  
Padre de Familia  
42574332  
Wilson Espinoza Juan W

## Matriz de consistencia

### MATRIZ DE CONSISTENCIA

Aplicación Web utilizando árboles de decisión para el rendimiento académico del curso de Comunicación - I.E. 20955-24 Jorge Guevara Mellado							
Problemas	Objetivos	Hipótesis	VARIABLES E INDICADORES				Metodología
Problema general	Objetivo general	Hipótesis general	VARIABLE DEPENDIENTE		VARIABLE INDEPENDIENTE		1. Tipo de Investigación: Aplicada 2. Enfoque de investigación: Cuantitativo 3. Nivel de investigación: Correlacional 4. Diseño de la Investigación: No experimental de tipo transversal 5. Método de la Investigación: Deductivo 6. Población: 42 estudiantes de quinto y sexto grado de primaria 7. Muestra: 37 estudiantes de quinto y sexto grado 8. Técnica e instrumento de recolección de datos:
¿De qué manera se relaciona la aplicación web utilizando árboles de decisión con el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado?	Determinar de qué manera se relaciona la aplicación web utilizando árboles de decisión con el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.	La aplicación web utilizando árboles de decisión influye significativamente el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.	Rendimiento Académico		Aplicación Web		
Problemas específicos	Objetivos específicos	Hipótesis específicas	Dimensiones	Indicadores	Desempeño / Ítem	Niveles	
¿De qué manera se relaciona la aplicación web utilizando árboles de decisión con la comunicación oral de la lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado?	Determinar de qué manera se relaciona la aplicación web utilizando árboles de decisión con la comunicación oral de la lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.	La aplicación web utilizando árboles de decisión influye significativamente con la comunicación oral de la lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.	Se comunica oralmente en su lengua materna	Infiere e interpreta información del texto oral	Nivel Destacado Nivel Esperado En proceso En inicio	AD A B C	
				Adecua, organiza y desarrolla las ideas de forma coherente y cohesionada	Nivel Destacado Nivel Esperado En proceso En inicio	AD A B C	
				Interactúa estratégicamente con distintos interlocutores	Nivel Destacado Nivel Esperado En proceso En inicio	AD A B C	
				Obtiene información	Nivel Destacado	AD	

			Lee diversos tipos de textos escritos en su lengua materna	del texto escrito	Nivel Esperado En proceso En inicio	A B C	Análisis documental, rubrica de evaluación y registro de notas 9. Método análisis de datos: Estadística descriptiva 10. Técnica para el procesamiento de datos: Programa estadístico SPSS versión 22
				Infiere e interpreta información del texto	Nivel Destacado Nivel Esperado En proceso En inicio	AD A B C	
¿De qué manera se relaciona la aplicación web utilizando árboles de decisión con leer diversos tipos de textos escritos en su lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado?	Determinar de qué manera se relaciona la aplicación web utilizando árboles de decisión con leer diversos tipos de textos escritos en su lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.	La aplicación web utilizando árboles de decisión influye significativamente con leer diversos tipos de textos escritos en su lengua materna en el rendimiento académico del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.	Escribe diversos tipos de textos en su lengua materna	Reflexiona y evalúa la forma, el contenido y contexto del texto	Nivel Destacado Nivel Esperado En proceso En inicio	AD A B C	
				Adecua el texto a la situación comunicativa	Nivel Destacado Nivel Esperado En proceso En inicio	AD A B C	
¿De qué manera se relaciona la aplicación web utilizando árboles de decisión con escribir diversos tipos de textos en su lengua materna en el rendimiento académico de los estudiantes del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado?	Determinar de qué manera se relaciona la aplicación web utilizando árboles de decisión para escribir diversos tipos de textos en su lengua materna en el rendimiento académico de los estudiantes del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.	La aplicación web utilizando árboles de decisión influye significativamente con escribir diversos tipos de textos en su lengua materna en el rendimiento académico de los estudiantes del curso de comunicación en la I.E. 20955-24 Jorge Guevara Mellado.		Organiza y desarrolla las ideas de forma coherente y cohesionada	Nivel Destacado Nivel Esperado En proceso En inicio	AD A B C	
				Utiliza convenciones del lenguaje escrito de forma pertinente	Nivel Destacado Nivel Esperado En proceso En inicio	AD A B C	

## Autorización de la aplicación del Instrumento

SOLICITO: APLICACIÓN  
DEL INSTRUMENTO

DIRECTORA DE LA INSTITUCIÓN EDUCATIVA N° 20955-24 "JORGE GUEVARA MELLADO"

Mg. Sue Marilyn Ocaris Cisneros

Tenemos el agrado de dirigirnos a usted, para saludarla cordialmente y a la vez manifestarle lo siguiente:

Por medio de la presente solicitamos a Usted, autorizar el recojo de datos mediante una rúbrica de evaluación la cual es el instrumento de investigación de las estudiantes **Mariana Ramón Vitteri** y **Becky Kazumy Velásquez Espinoza**, correspondiente al Proyecto de investigación sobre **"APLICACIÓN WEB UTILIZANDO ÁRBOLES DE DECISIÓN PARA EL RENDIMIENTO ACADÉMICO DEL CURSO DE COMUNICACIÓN - I.E. 20955-24 JORGE GUEVARA MELLADO"**.

Nuestro objetivo es recoger la información de los estudiantes con respecto a notas y asistencias de los últimos 3 años para analizar el impacto de nuestra aplicación web en el rendimiento académico en el curso de Comunicación. Solicitamos su autorización para acceder a estos datos y su apoyo para coordinar con los docentes responsables. Garantizamos que toda la información será manejada con la máxima confidencialidad y utilizada únicamente con fines académicos.

La recolección de datos se llevará a cabo durante el mes de abril y se estima que tomará aproximadamente una semana. Agradecemos de antemano su apoyo y colaboración para coordinar los tiempos y espacios necesarios sin interferir en las clases. Nos comprometemos a compartir los resultados obtenidos con la institución para contribuir al mejoramiento continuo de la calidad educativa.

Agradecemos su atención y quedamos a su disposición para cualquier consulta o aclaración adicional que requiera.

Vitarte 10 de abril del 2024

Atentamente,



Mariana Ramón Vitteri

Estudiante



Director  
I.E. N° 20955-24 "J.G.M."

Mg. Sue Marilyn Ocaris Cisneros

Directora



Becky Velásquez Espinoza

Estudiante

## Base de datos en Excel y SPSS

Nº Estudiantes	2021			2022			2023		
	NL1	NL2	NL3	NL1	NL2	NL3	NL1	NL2	NL3
1	A	B	B	A	B	B	A	B	B
2	A	A	B	AD	A	A	AD	A	A
3	A	B	B	A	A	A	A	A	A
4	A	A	A	A	AD	A	A	AD	A
5	B	A	A	B	A	A	B	A	A
6	A	A	B	A	A	B	A	A	B
7	AD	A	A	AD	AD	A	AD	AD	A
8	A	A	C	A	A	C	A	A	C
9	A	A	A	A	A	A	A	A	A
10	A	B	B	A	B	B	A	B	B
11	C	B	B	C	B	C	C	B	C
12	A	A	B	A	A	A	A	A	A
13	AD	A	A	AD	AD	A	AD	AD	A
14	A	B	B	A	B	B	A	B	B
15	B	A	A	A	A	A	A	A	A
16	A	B	B	A	B	B	A	B	B
17	A	B	B	A	B	B	A	B	B
18	AD	A	A	AD	A	A	AD	A	A
19	B	A	A	B	A	A	B	A	A
20	B	A	A	B	A	A	A	A	A
21	B	B	A	B	B	A	B	B	A
22	B	A	B	B	A	B	A	A	B
23	A	A	A	AD	AD	A	AD	AD	A
24	A	A	B	A	A	C	A	A	C
25	A	A	B	A	A	A	A	A	A
26	C	B	B	C	B	C	C	B	C
27	A	A	A	A	A	A	A	A	A
28	A	B	B	A	B	B	A	B	B
29	B	A	A	B	A	A	A	A	A
30	A	B	B	A	B	B	A	B	B
31	A	B	B	A	A	A	A	A	A
32	B	B	A	B	A	A	B	A	A
33	A	B	B	A	A	B	A	A	B
34	AD	A	A	AD	AD	A	AD	AD	A
35	A	B	A	A	A	A	A	A	A
36	A	B	B	A	B	B	A	B	B
37	B	B	B	C	B	B	C	B	B

BASE DE DATOS SPSS.sav [Conjunto\_de\_datos1] - IBM SPSS Statistics Editor de datos

Archivo Editar Ver Datos Transformar Analizar Marketing directo Gráficos Utilidades Ventana Ayuda

26 :

	NL1_2021	NL2_2021	NL3_2021	NL1_2022	NL2_2022	NL3_2022	NL1_2023	NL2_2023	NL3_2023	NP_2024
1	3,00	2,00	2,00	3,00	2,00	2,00	3,00	2,00	2,00	3,00
2	3,00	3,00	2,00	4,00	3,00	3,00	4,00	3,00	3,00	4,00
3	3,00	2,00	2,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00
4	3,00	3,00	3,00	3,00	4,00	3,00	3,00	4,00	3,00	3,00
5	2,00	3,00	3,00	2,00	3,00	3,00	2,00	3,00	3,00	3,00
6	3,00	3,00	2,00	3,00	3,00	2,00	3,00	3,00	2,00	3,00
7	4,00	3,00	3,00	4,00	4,00	3,00	4,00	4,00	3,00	4,00
8	3,00	3,00	1,00	3,00	3,00	1,00	3,00	3,00	1,00	3,00
9	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00
10	3,00	2,00	2,00	3,00	2,00	2,00	3,00	2,00	2,00	3,00
11	1,00	2,00	2,00	1,00	2,00	1,00	1,00	2,00	1,00	2,00
12	3,00	3,00	2,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00
13	4,00	3,00	3,00	4,00	4,00	3,00	4,00	4,00	3,00	4,00
14	3,00	2,00	2,00	3,00	2,00	2,00	3,00	2,00	2,00	3,00
15	2,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00
16	3,00	2,00	2,00	3,00	2,00	2,00	3,00	2,00	2,00	3,00
17	3,00	2,00	2,00	3,00	2,00	2,00	3,00	2,00	2,00	3,00
18	4,00	3,00	3,00	4,00	3,00	3,00	3,00	4,00	3,00	3,00
19	2,00	3,00	3,00	2,00	3,00	3,00	2,00	3,00	3,00	3,00
20	2,00	3,00	3,00	2,00	3,00	3,00	3,00	3,00	3,00	4,00
21	2,00	2,00	3,00	2,00	2,00	3,00	2,00	2,00	3,00	2,00
22	2,00	3,00	2,00	2,00	3,00	2,00	3,00	3,00	2,00	3,00
23	3,00	3,00	3,00	4,00	4,00	3,00	4,00	4,00	3,00	4,00
24	3,00	3,00	2,00	3,00	3,00	1,00	3,00	3,00	1,00	3,00
25	3,00	3,00	2,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00
26	1,00	2,00	2,00	1,00	2,00	1,00	1,00	2,00	1,00	2,00
27	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00

Vista de datos Vista de variables

# Resultados SPSS

\*Resultados.spv [Documento1] - IBM SPSS Statistics Visor

Archivo Editar Ver Datos Transformar Insertar Formato Analizar Marketing directo Gráficos Utilidades Ventana Ayuda

Resultado

- Frecuencias
  - Estadísticos

```

FRECUENCIAS VARIABLES=NL1_2021 NL2_2021 NL3_2021 NL1_2022 NL2_2022 NL3_2022 NL1_2023 NL2_2023 NL3_2023
/STATISTICS=STDDEV VARIANCE MINIMUM MAXIMUM SEMEAN MEAN SKEWNESS SESKEW
/HISTOGRAM NORMAL
/ORDER=ANALYSIS.
    
```

### Frecuencias

Estadísticos

		D1 - 2021	D2 - 2021	D3 - 2021	D1 - 2022	D2 - 2022	D3 - 2022	D1 - 2023	D2 - 2023	D3 - 2023
N	Válido	37	37	37	37	37	37	37	37	37
	Perdidos	0	0	0	0	0	0	0	0	0
Media		2,7568	2,5405	2,4054	2,8108	2,8108	2,4595	2,8919	2,8108	2,4595
Error estándar de la media		,11884	,08306	,09054	,13332	,10848	,11360	,12725	,10848	,11360
Desviación estándar		,72286	,50523	,55073	,81096	,65988	,69100	,77401	,65988	,69100
Varianza		,523	,255	,303	,658	,435	,477	,599	,435	,477
Asimetría		-,524	-,170	-,131	-,623	,219	-,915	-,947	,219	-,915
Error estándar de asimetría		,388	,388	,388	,388	,388	,388	,388	,388	,388
Mínimo		1,00	2,00	1,00	1,00	2,00	1,00	2,00	1,00	1,00
Máximo		4,00	3,00	3,00	4,00	4,00	3,00	4,00	4,00	3,00

### Tabla de frecuencia

D1 - 2021

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	En inicio (C)	2	5,4	5,4	5,4
	En proceso (B)	9	24,3	24,3	29,7
	Logro esperado (A)	22	59,5	59,5	89,2
	Logro destacado (AD)	4	10,8	10,8	100,0
	Total	37	100,0	100,0	

D2 - 2021

\*Resultados.spv [Documento1] - IBM SPSS Statistics Visor

Archivo Editar Ver Datos Transformar Insertar Formato Analizar Marketing directo Gráficos Utilidades Ventana Ayu

Resultado

- Fiabilidad
  - Estadísticos

### Fiabilidad

Escala: ALL VARIABLES

Resumen de procesamiento de casos

		N	%
Casos	Válido	37	100,0
	Excluido <sup>a</sup>	0	,0
	Total	37	100,0

a. La eliminación por lista se basa en todas las variables del procedimiento.

Estadísticas de fiabilidad

	Alfa de Cronbach	N de elementos
	,904	9

```

EXAMINE VARIABLES=NL1_2021 NL2_2021 NL3_2021
/PLOT NPFPLOT
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
    
```

### Explorar

Resumen de procesamiento de casos

	Casos					
	Válido		Perdidos		Total	
	N	Porcentaje	N	Porcentaje	N	Porcentaje
D1 - 2021	37	100,0%	0	0,0%	37	100,0%
D2 - 2021	37	100,0%	0	0,0%	37	100,0%

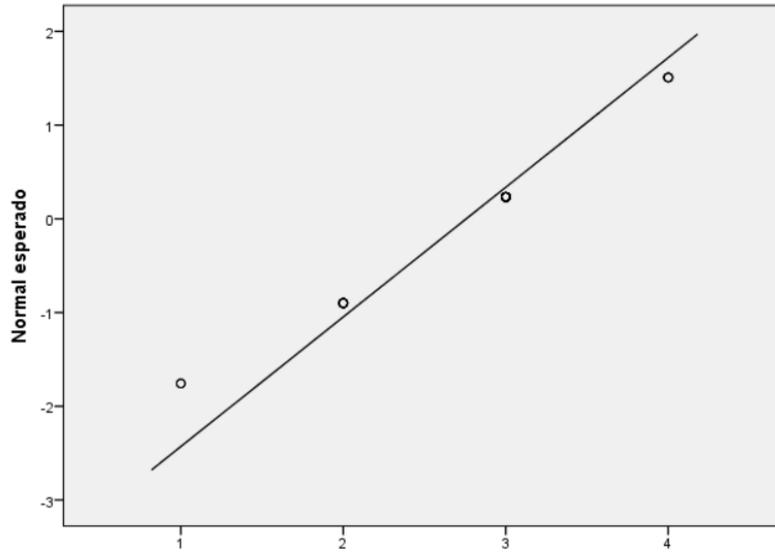
**Pruebas de normalidad**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
D1 - 2021	,334	37	,000	,816	37	,000
D2 - 2021	,359	37	,000	,635	37	,000
D3 - 2021	,337	37	,000	,709	37	,000

a. Corrección de significación de Lilliefors

**D1 - 2021**

**Gráfico Q-Q normal de D1 - 2021**



IBM SPSS Statistics Viewer interface showing a tree view on the left and a non-parametric correlations table on the right.

**Correlaciones no paramétricas**

			Correlaciones									
			D1 - 2021	D2 - 2021	D3 - 2021	D1 - 2022	D2 - 2022	D3 - 2022	D1 - 2023	D2 - 2023	D3 - 2023	NP_2024
Rho de Spearman	D1 - 2021	Coefficiente de correlación	1,000	,167	-,014	,916**	,361*	,154	,821**	,361*	,154	,562**
		Sig. (bilateral)	.	,322	,934	,000	,028	,364	,000	,028	,364	,000
		N	37	37	37	37	37	37	37	37	37	37
D2 - 2021		Coefficiente de correlación	,167	1,000	,428**	,289	,750**	,448**	,431**	,750**	,448**	,436**
		Sig. (bilateral)	,322	.	,008	,083	,000	,005	,008	,000	,005	,007
		N	37	37	37	37	37	37	37	37	37	37
D3 - 2021		Coefficiente de correlación	-,014	,428**	1,000	,078	,547**	,761**	,154	,547**	,761**	,370*
		Sig. (bilateral)	,934	,008	.	,646	,000	,000	,364	,000	,000	,024
		N	37	37	37	37	37	37	37	37	37	37
D1 - 2022		Coefficiente de correlación	,916**	,289	,078	1,000	,456**	,264	,920**	,456**	,264	,677**
		Sig. (bilateral)	,000	,083	,646	.	,005	,115	,000	,005	,115	,000
		N	37	37	37	37	37	37	37	37	37	37
D2 - 2022		Coefficiente de correlación	,361*	,750**	,547**	,456**	1,000	,634**	,544**	1,000**	,634**	,651**
		Sig. (bilateral)	,028	,000	,000	,005	.	,000	,001	.	,000	,000
		N	37	37	37	37	37	37	37	37	37	37
D3 - 2022		Coefficiente de correlación	,154	,448**	,761**	,264	,634**	1,000	,315	,634**	1,000**	,476**
		Sig. (bilateral)	,364	,005	,000	,115	,000	.	,058	,000	.	,003
		N	37	37	37	37	37	37	37	37	37	37
D1 - 2023		Coefficiente de correlación	,821**	,431**	,154	,920**	,544**	,315	1,000	,544**	,315	,786**
		Sig. (bilateral)	,000	,008	,364	,000	,001	,058	.	,001	,058	,000
		N	37	37	37	37	37	37	37	37	37	37
D2 - 2023		Coefficiente de correlación	,361*	,750**	,547**	,456**	1,000	,634**	,544**	1,000	,634**	,651**
		Sig. (bilateral)	,028	,000	,000	,005	.	,000	,001	.	,000	,000
		N	37	37	37	37	37	37	37	37	37	37

## Carta de conformidad de Investigación

**“Año del Bicentenario, de la consolidación de nuestra Independencia, y de la conmemoración de las heroicas batallas de Junín y Ayacucho”.**

### CARTA DE CONFORMIDAD DE INVESTIGACIÓN DE TESIS

San Pedro de Mama, 11 de julio 2024

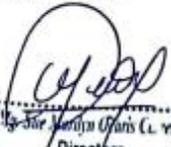
#### ESCUELA PROFESIONAL DE INGENIERIA DE SISTEMAS

Por la presente, la directora de la I.E. 20955-24 “Jorge Guevara Mellado” certifica que las estudiantes **Ramón Vitteri Mariana** identificada con DNI N° 70570204 y **Velásquez Espinoza Becky Kazumy** identificada con DNI N° 71403770, de la carrera de ingeniería de sistemas de la **Universidad César Vallejo**, han culminado satisfactoriamente su investigación de tesis titulada **“Aplicación Web utilizando árboles de decisión para el rendimiento académico del curso de Comunicación - I.E. 20955-24 Jorge Guevara Mellado”** en nuestra institución, desarrollada de manera adecuada y cumplido con todos los objetivos planteados.

Las estudiantes demostraron compromiso y profesionalismo a lo largo del proyecto, contribuyendo de manera significativa al avance del conocimiento de nuestro entorno educativo. Los resultados obtenidos han sido de gran importancia para nuestra institución.

Agradecemos a **Ramón Vitteri Mariana** y **Velásquez Espinoza Becky Kazumy**, así como a la **Universidad César Vallejo**, por la oportunidad de colaborar en esta investigación.

Atentamente,


Sue Marilyn Ocaris Cisneros  
Directora  
I.E.I. N° 20955-24 "J.G.M."

Mag. Sue Marilyn Ocaris Cisneros

Directora

- **Artículo de Resultados**

## *Decision tree-based artificial intelligence for academic performance in an educational institution: a case study.*

Velasquez Espinoza Becky Kazumy<sup>1</sup>, Ramon Vitteri Mariana<sup>2</sup>, Giancarlo Sánchez Atuncar<sup>3</sup>

School of Systems Engineering, Universidad Cesar Vallejo, Lima, Peru<sup>1,2,3</sup>

[bvelasquez@ucvvirtual.edu.pe](mailto:bvelasquez@ucvvirtual.edu.pe)

[rmaronv@ucvvirtual.edu.pe](mailto:rmaronv@ucvvirtual.edu.pe)

[gsanchezat@ucvvirtual.edu.pe](mailto:gsanchezat@ucvvirtual.edu.pe)

**Abstract.** The objective of this study was to improve students' academic performance by predicting their performance in advance, thus enabling the development of effective educational strategies. A decision tree-based artificial intelligence model was implemented using the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology to guide the development and implementation of data analysis, which is composed of six main phases: business understanding, data understanding, data preparation, modeling, evaluation, and deployment [10]. DT was selected due to its ability to handle both numerical and categorical data, as well as its ease of interpretation. A previous study by [15] used the Decision Trees algorithm, proving to be effective with an accuracy of over 76.9% in predicting students' academic performance. Using the `sklearn` library, we proceeded to segment the data into sets intended for training and testing with the purpose of analyzing the performance of the model with metrics such as accuracy and area under the ROC curve. In addition, a cross-validation was performed, obtaining a mean precision of 0.975 and a standard deviation of 0.05, which demonstrates high precision and low variability between folds.

**Keywords.** Artificial intelligence, Decision trees, Academic performance, Machine learning, Predictive model

I INTRODUCTION

interpretations of the determinants in different contexts. [5]. These models, which replicate the human decision-making process by creating branching structures of logical rules, offer a way to identify patterns and establish connections to address challenges through data analysis, with the primary goal of forecasting future trends. [6] In contrast to more complex methods such as with other supervised learning models like, logistic regression and support vector machines (SVMs), DTs are highly interpretable and can effectively handle nonlinear relationships between variables without requiring additional adjustments. [7]. In this regard, decision trees achieve an accuracy of 82.87%, in contrast to logistic regression which achieves 74.4% and support vector machines (SVMs) which obtain 75.2% in estimating academic performance. [8].

Academic achievement is essential in defining educational success, as it is based on students' performance in their learning activities and assessments. [9]. This aspect is crucial, as it impacts both the personal development of students and the success of educational institutions by reflecting the relationship between the work of teachers and students, as well as the educational landscape as a whole. [10]. The fundamental purpose of academic performance is to reveal the factors that affect the outcomes of the educational process. Therefore,

- **Articulo Literario**

## *Comparative study of artificial intelligence models for improving academic performance: A systematic review of the literature.*

Velasquez Espinoza Becky Kazumy<sup>1</sup>, Ramon Vitteri Mariana<sup>2</sup>, Giancarlo Sánchez Atuncar<sup>3</sup>

School of Systems Engineering, Universidad Cesar Vallejo, Lima, Peru<sup>1,2,3</sup>

[bvelasquez@ucvvirtual.edu.pe](mailto:bvelasquez@ucvvirtual.edu.pe)

[rmaronv@ucvvirtual.edu.pe](mailto:rmaronv@ucvvirtual.edu.pe)

[gsanchezat@ucvvirtual.edu.pe](mailto:gsanchezat@ucvvirtual.edu.pe)

**Abstract.** The objective of this research was to compare different models of artificial intelligence to improve academic performance, evaluating studies published between 2019-2024 in five databases: Science Direct, Taylor & Francis, IEEE Xplore, Scopus and Springer. The PRISMA methodology was employed, selecting 30 articles, with the year of highest publication being 2021. In addition, all the selected publications were written in English language. Finally, it was found that the most effective artificial intelligence models for improving academic performance include neural networks, deep learning algorithms, support vector machines and decision trees. The results showed a significant improvement in students' academic performance when implementing these models. In addition, it was highlighted that it is crucial to consider external factors such as educational context, type of students, subject matter, and availability of technological resources before selecting and implementing an artificial intelligence model as an educational tool.

**Keywords.** Artificial intelligence, Academic performance, Deep learning, Machine learning, Machine learning.

I INTRODUCTION

In recent years, education has undergone a significant transformation by adopting technological tools and innovative methodologies to improve academic performance. [1]. The use of predictive analytics techniques, such as decision trees, has been highlighted as a promising approach to understand and optimize student performance in educational settings. [2]. As the world has embraced digital transformation, educational institutions are exploring new ways to leverage data and

In addition to DTs, models such as random forests, neural networks, support vector machines (SVMs), and logistic regression will be considered. Each offers unique approaches to analyzing data and predicting academic outcomes. For example, random forests combine multiple decision trees to improve accuracy and reduce overfitting [5], while neural networks are highly adaptive and effective in complex tasks such as natural language processing and computer vision. [6].

Academic performance is critical in today's education, especially in school contexts that seek to optimize available resources for student success. The implementation of decision trees offers a systematic approach to identify key variables that directly impact student academic performance. [7]. It can provide a comprehensive view that allows educators and decision makers to make informed decisions to improve student outcomes. [8].

It is essential to understand that the use of decision trees in education goes beyond simple data collection; it involves the ability to transform information into meaningful actions that can positively influence student learning and development. [9]. Just as companies adopt innovative strategies to compete in a digital environment, educational institutions are exploring new ways to optimize available resources to promote academic success. This approach focuses on leveraging advanced analytical tools to understand and optimize the key factors that influence student success.