

**UNIVERSIDAD NACIONAL JORGE BASADRE GROHMANN**

**Facultad de Ingeniería**

Escuela Profesional de Ingeniería en Informática y Sistemas

**DETERMINACIÓN DEL EFECTO DEL USO DE TICs EN  
EL RENDIMIENTO ACADÉMICO DE LOS  
ESTUDIANTES DE LA FACULTAD DE CIENCIAS DE  
LA SALUD EN TIEMPOS DE PANDEMIA,  
UTILIZANDO MINERÍA DE DATOS EN LA  
UNIVERSIDAD NACIONAL JORGE  
BASADRE GROHMANN, 2020**

**TESIS**

Presentada por:

Bach. Arie Belén Yufra Suárez

Para optar el Título Profesional de:

**INGENIERO EN INFORMÁTICA Y SISTEMAS**

TACNA – PERÚ

2025

**UNIVERSIDAD NACIONAL JORGE BASADRE GROHMANN**  
**FACULTAD DE INGENIERÍA**  
**ESCUELA PROFESIONAL DE INGENIERÍA EN INFORMÁTICA Y SISTEMAS**

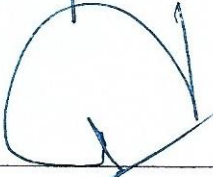
**“DETERMINACIÓN DEL EFECTO DEL USO DE TICS EN EL  
RENDIMIENTO ACADÉMICO DE LOS ESTUDIANTES DE LA FACULTAD  
DE CIENCIAS DE LA SALUD EN TIEMPOS DE PANDEMIA. UTILIZANDO  
MINERÍA DE DATOS EN LA UNIVERSIDAD NACIONAL JORGE  
BASADRE GROHMANN, 2020”**

TESIS PRESENTADA Y APROBADA EL 04 DE JUNIO DEL 2024


ESTANDO EL JURADO CALIFICADOR INTEGRADO POR:

Presidente :   
\_\_\_\_\_


Mtro. Hugo Manuel Barraza Vizcarra

Secretario :   
\_\_\_\_\_

Mag. Luis Johnson Paúl Mori Sosa Tejada

Vocal :   
\_\_\_\_\_

Dr. Edgar Aurelio Taya Acosta

Asesor :   
\_\_\_\_\_

Dr. Edgar Aurelio Taya Acosta

## CERTIFICADO DE SIMILITUD

Yo Edgar Aurelio Taya Acosta, en mi condición de asesor acreditado por la Resolución de Facultad N° 08502-2024-FAIN/UNIBG de la tesis:

“DETERMINACIÓN DEL EFECTO DEL USO DE LAS TICs EN EL RENDIMIENTO ACADÉMICO DE LOS ESTUDIANTES DE LA FACULTAD DE CIENCIAS DE LA SALUD EN TIEMPOS DE PANDEMIA, UTILIZANDO MINERÍA DE DATOS EN LA JORGE BASADRE GROHMANN, 2020”.

Presentado por el Bachiller Arieé Belén Yufra Suárez para optar el título profesional de Ingeniero en Informática y Sistemas.

Habiendo cumplido con lo establecido en el reglamento de originalidad y de similitud de trabajo de investigación y producción intelectual, considerando que según la revisión, evaluación y análisis realizado a través del **software de similitud textual TURNITIN**, cuenta con el **nivel de similitud permitido cuyo porcentaje es 7%**. Por lo que, **CERTIFICO LA SIMILARIDAD** de la tesis enunciado líneas de arriba, la cual está expedita para continuar con los trámites para la obtención del Título Profesional según corresponda consiguientemente la publicación en el repositorio institucional.

---

Dr. Edgar Aurelio Taya Acosta  
DNI: 00509033  
ASESOR



---

Bach. Arieé Belén Yufra Suárez  
DNI: 71227664  
TESISTA



## *Dedicatoria*

*Dedico esta tesis a mis padres y a  
mis hermanas, quienes han estado  
conmigo superando cualquier  
adversidad.*

## *Agradecimientos*

*A Dios, mis padres y mis hermanas por confiar en mí y apoyarme en toda mi etapa universitaria.*

*A mi casa de estudios, la Universidad Nacional Jorge Basadre Grohmann, a la Escuela Profesional de Ingeniería de Informática y Sistemas, al asesor de mi tesis, Dr. Edgar Taya Acosta por su apoyo y paciencia.*

*Finalmente, a los docentes que tuve a lo largo de toda la carrera universitaria que, con su apoyo y enseñanzas forman la base de mi vida profesional.*

## CONTENIDO

CONTENIDO	iv
ÍNDICE DE TABLAS	vi
ÍNDICE DE FIGURAS	vii
RESUMEN	viii
ABSTRACT	ix
INTRODUCCIÓN	1
PLANTEAMIENTO DE LA INVESTIGACIÓN	3
1.1. Descripción del Problema	3
1.1.1. Antecedentes del Problema	3
1.1.2. Problemática de la investigación	6
1.2. Formulación del problema	7
1.3. Justificación e importancia	7
1.4. Alcances y limitaciones	9
1.5. Objetivos	9
1.5.1. Objetivo general	10
1.5.2. Objetivos específicos	10
1.6. Hipótesis	10
1.6.1. Hipótesis general	10
1.6.2. Subhipótesis	11
MARCO TEÓRICO	13
2.1. Antecedentes del estudio	13
2.2. Bases teóricas	14
2.2.1. Tecnologías de la información y Comunicación (TIC)	14
2.2.2. Rendimiento Académico	14
2.2.3. Minería de datos	15

2.2.4. Machine learning	15
2.2.5. Aprendizaje supervisado	15
2.2.6. EDM (Educational Data Mining)	15
2.3. Definición de términos	16
2.3.1. Coronavirus COVID-19	16
2.3.2. Pandemia	16
2.3.3. Proceso KDD para la obtención del conocimiento	16
2.3.4. Aprendizaje supervisado y no supervisado	17
2.3.5. Árboles de decisión (Decision Tree DT)	17
2.3.6. Random Forest	19
2.3.7. Matriz de confusión	19
MARCO METODOLÓGICO	21
3.1. Tipo y diseño de investigación	21
3.2. Población y muestra	22
3.3. Operacionalización de variables	22
3.4. Técnicas e instrumentos para recolección de datos	24
3.5. Procesamiento y análisis de datos	25
RESULTADOS Y DISCUSIÓN	27
4.1. Resultados de la aplicación del marco de trabajo KDD	27
4.2. Resultados de descriptivos de las variables	30
4.3. Resultados de los modelos aplicados	34
4.4. Resultados a nivel inferencial	36
4.5. Discusión	38
CONCLUSIONES	41
RECOMENDACIONES	43
REFERENCIAS BIBLIOGRAFICAS	45
ANEXOS	49

## ÍNDICE DE TABLAS

Tabla 1 Matriz de confusión	19
Tabla 2. Operacionalización de variables	23
Tabla 3. Matriz de confusión Árbol de decisión	35
Tabla 4. Matriz de confusión Random Forest	35

## ÍNDICE DE FIGURAS

Figura 1. Proceso de KDD en la minería de datos.	17
Figura 2. La función booleana $Y = X1 \wedge X2 \vee X3$ y su representación en DT	18
Figura 3. Modelo propuesto para la investigación	21
Figura 4. Conjunto de datos Inicial	28
Figura 5. Reemplazo de valores faltantes por la media	28
Figura 6. Notas finales anonimizadas de todos los alumnos	29
Figura 7. Variables más importantes	29
Figura 8. Consumo de streaming por escuela	30
Figura 9. Promedio de notas por escuela	31
Figura 10. Frecuencia de uso de aplicaciones de mensajería en horas	32
Figura 11. Uso de redes sociales por escuela y edad	33

## RESUMEN

El presente trabajo de investigación se enfoca en determinar el efecto de uso de las Tecnologías de la Información y la Comunicación (TIC) en el rendimiento académico de los estudiantes de la Facultad de Ciencias de la Salud durante el año 2020, en el contexto de la pandemia. Por lo cual, se analizaron variables como las horas en redes sociales, plataformas de streaming, apps de mensajería, y notas finales de curso.

Para el análisis de datos se utilizó un cuestionario estructurado como instrumento de recolección y se aplicaron técnicas de minería de datos, específicamente los algoritmos Random Forest y Árbol de Decisión, que permitieron identificar patrones significativos en el uso de TIC y su influencia en el rendimiento académico.

Finalmente, los resultados evidenciaron que el uso de las TIC si tiene un efecto significativo en el rendimiento académico, aunque con variaciones según el tipo de herramienta utilizada, siendo las aplicaciones de mensajería las de mayor impacto positivo.

**Palabras clave:** Tecnologías de la Información y la Comunicación, Rendimiento académico, Minería de datos, Random Forest, Árbol de decisión, Aplicaciones de mensajería.

## **ABSTRACT**

This research focuses on determining the effect of the use of Information and Communication Technologies (ICT) on the academic performance of students from the Faculty of Health Sciences during the year 2020, in the context of the COVID-19 pandemic. Variables such as hours spent on social networks, streaming platforms, messaging apps, and final course grades were analyzed.

For data analysis, a structured questionnaire was used as the data collection instrument, and data mining techniques were applied, specifically the Random Forest and Decision Tree algorithms. These allowed the identification of significant patterns in the use of ICT and their influence on academic performance.

The results show that the use of ICT has a significant effect on academic performance, although the impact varies depending on the type of tool used, with messaging applications having the greatest positive influence.

**Keywords:** Information and Communication Technologies, Academic performance, Data mining, Random Forest, Decision Tree, Messaging applications.

## INTRODUCCIÓN

La pandemia de COVID-19, iniciada a fines de 2019, ha generado cambios significativos en el ámbito educativo, obligando a las instituciones a adoptar rápidamente la enseñanza virtual. En este contexto, la Facultad de Ciencias de la Salud de la Universidad Nacional Jorge Basadre Grohmann, en Tacna, Perú, tuvo que adaptarse a esta nueva realidad utilizando diversas herramientas tecnológicas.

Este estudio tiene como objetivo determinar el efecto del uso de las Tecnologías de la Información y la Comunicación (TIC) en el rendimiento académico de los estudiantes durante el año 2020. Utilizando técnicas de minería de datos, se analizarán grandes volúmenes de información para descubrir patrones relevantes que ayuden a comprender esta relación.

El problema de investigación se centra en cómo las herramientas tecnológicas utilizadas durante la pandemia han influido en los resultados académicos de los estudiantes. La pregunta principal es: ¿Cómo ha afectado el uso intensivo de las TIC el rendimiento académico de los estudiantes de la Facultad de Ciencias de la Salud durante el 2020?

Este estudio es relevante porque puede ofrecer una visión clara sobre los desafíos y oportunidades de las TIC en la educación, especialmente en tiempos de

crisis. Los hallazgos pueden contribuir al desarrollo de estrategias educativas más efectivas y equitativas.

En las siguientes secciones se detallará la descripción del problema, los antecedentes, la formulación del problema, la justificación, los objetivos, las hipótesis, el marco teórico, el diseño metodológico, y finalmente, los resultados y discusión, conclusiones, recomendaciones y referencias.

# CAPÍTULO I

## PLANTEAMIENTO DE LA INVESTIGACIÓN

### 1.1. Descripción del Problema

En el contexto de la pandemia de la COVID-19, la educación superior ha experimentado una transformación significativa en la forma en que se imparten las clases. La Universidad Nacional Jorge Basadre Grohmann, al igual que muchas otras instituciones educativas en todo el mundo, se vio obligada a adoptar la enseñanza en línea y el uso de Tecnologías de la Información y Comunicación (TICs) como una medida de emergencia para garantizar la continuidad de la educación.

Este cambio abrupto plantea una serie de preguntas críticas sobre cómo afecta el uso de TIC al rendimiento académico de los estudiantes, especialmente en una facultad específica, en este caso, la Facultad de Ciencias de la Salud. La minería de datos, como una herramienta analítica, ofrece la posibilidad de abordar estas preguntas de manera sistemática y basada en evidencia.

#### 1.1.1. Antecedentes del Problema

Para una mejor comprensión del problema, se resume una serie de investigaciones relacionadas con la problemática abordada en este estudio:

(Dhiman, 2022) en su investigación titulada “Use and Impact of Social Media on Academic Performance of Kurukshetra University Students: A Case Study” nos comenta que, con el aumento de la popularidad de las redes sociales, se preguntan si las calificaciones de los estudiantes no se verán afectadas por el tiempo que pasan en estos sitios. Para evaluar el uso de las redes sociales y su impacto en el rendimiento académico, utilizaron un cuestionario como recolección de datos aplicado a 120 estudiantes. Teniendo como resultados, con un nivel significativo de 0.01, si le dedican más tiempo a las redes sociales tendrá un rendimiento académico bajo.

(Xu et al., 2019) en su investigación titulada “Prediction of academic performance associated with internet usage behaviors using machine learning algorithms” revela la asociación que hay entre el uso de internet y el rendimiento académico y predecir el rendimiento académico de un total de 4000 estudiantes. Para predecir el rendimiento académico se dividió en tres grupos: Grupo 1 comprende las características de la categoría de duración en línea; Grupo 2 contiene funciones de categorías de duración en línea y volumen de Internet. Además de las funciones del Grupo 1 anterior, hay tres funciones de tráfico de Internet; Grupo 3 se compone de las tres categorías de funciones, en las que se suman la frecuencia de conexión móvil, la frecuencia de conexión de PC y la frecuencia de conexión total en función del Grupo

2. Se utilizaron tres algoritmos de aprendizaje automático de árbol de decisión, red neuronal y máquina de vectores de soporte.

(Ullah et al., 2019) en su investigación titulada “Impact of ICT on Students’ Academic Performance: Applying Association Rule Mining and Structured Equation Modeling” indica que examinó el efecto de las TIC en el rendimiento académico de estudiantes de diferentes universidades privadas de Chittagong, Bangladesh. Para evaluar el efecto utilizaron estadística y minería de datos, como la minería de reglas de asociación, los algoritmos A priori y eclat R usando un paquete llamado regla; teniendo como resultado que en las pruebas estadísticas de significancia y valor r cuadrado su principal predictor es la “adicción a las TIC” y en la minería de datos de asociación también es un predictor significativo.

(Asokan, 2019) en su investigación titulada “Internet Addiction among Medical Students and Its Impact on Academic Performance: An Indian Study” tuvo como objetivo determinar la prevalencia de la adicción a internet entre los estudiantes de medicina y su efecto en el rendimiento académico de una facultad de medicina. Para evaluar el análisis utilizaron estadística descriptiva que incluye correlación y regresión lineal. Se utilizó un cuestionario brindado a 381 estudiantes, 294 mujeres y 107 hombres. Se obtuvo como resultado que, el 61.4% tiene adicción a internet, entre ellos 149 estudiantes (63.7%) tenían adicción leve, 83 estudiantes (35.5%) moderada y 2

estudiantes (0.8%) leve, entonces la adicción a internet tuvo un impacto negativo en su rendimiento académico.

### **1.1.2. Problemática de la investigación**

En el contexto de la pandemia de la COVID-19, la educación superior ha experimentado una transformación significativa en la forma en que se imparten las clases. La Universidad Nacional Jorge Basadre Grohmann, al igual que muchas otras instituciones educativas en todo el mundo, se vio obligada a adoptar por la enseñanza en línea y el uso de Tecnologías de la Información y Comunicación (TICs) como una medida de emergencia para garantizar la continuidad de la educación.

Este cambio abrupto plantea una serie de preguntas críticas sobre cómo afecta el uso de TICs al rendimiento académico de los estudiantes, especialmente en una facultad específica, en este caso, la Facultad de Ciencias de la Salud. La minería de datos, como una herramienta analítica, ofrece la posibilidad de abordar estas preguntas de manera sistemática y basada en evidencia.

## **1.2. Formulación del problema**

Ya planteada la problemática de la presente investigación, se formula el problema general a través de la siguiente pregunta:

### **Problema general**

¿Cuál es el efecto del uso de las TICs en el rendimiento académico de los estudiantes de la Facultad de Ciencias de la Salud en tiempos de pandemia en la Universidad Nacional Jorge Basadre Grohmann, 2020?

### **Problemas específicos**

Para dar respuesta a la pregunta anterior, se formularon los siguientes problemas específicos:

- ¿Cuál es el efecto del uso de las redes sociales en el rendimiento académico?
- ¿Cuál es el efecto del consumo de streaming en el rendimiento académico?
- ¿Cuál es el efecto del uso de aplicaciones de mensajería en el rendimiento académico?

## **1.3. Justificación e importancia**

La pandemia de COVID-19 obligó a las instituciones educativas a cambiar rápidamente a un formato de aprendizaje en línea. Las TIC se convirtieron en herramientas esenciales para facilitar este cambio. Investigar cómo se utilizan estas

herramientas puede proporcionar información valiosa sobre cómo mejorar la educación en línea.

El rendimiento académico es un indicador clave del éxito educativo. Comprender cómo las TIC afectan este rendimiento puede ayudar a las instituciones a implementar estrategias más efectivas para mejorar los resultados de los estudiantes.

La pandemia ha resaltado las desigualdades existentes en términos de acceso a la tecnología y la capacidad para utilizarla eficazmente. Investigar estas cuestiones puede ayudar a identificar y abordar estas brechas.

Las TIC seguirán desempeñando un papel crucial en la educación y en el mundo laboral. Entender cómo se utilizan en un contexto educativo puede ayudar a preparar mejor a los estudiantes para el futuro.

La investigación permite una mejora continua, proporcionando datos y hallazgos que pueden informar las decisiones y políticas futuras. Una investigación de este tipo es vital para entender y mejorar la educación en tiempos de crisis y más allá.

## **1.4. Alcances y limitaciones**

### **Alcances**

La presente investigación se encuentra limitada por la disponibilidad de tiempo de los estudiantes de la Facultad de Medicina, lo cual afecta la cantidad de participantes en la muestra.

### **Limitaciones**

Se realizará la investigación solo con estudiantes de la Facultad de Ciencia de la Salud durante el año de ingreso a la universidad 2020. Para obtener información se utilizará un cuestionario dirigido a los estudiantes que ahora cursan ciclos más avanzados

## **1.5. Objetivos**

En esta sección se define el propósito de la presente investigación. Asimismo, se plantean objetivos específicos, los cuales nos permitirán alcanzar el objetivo general.

### **1.5.1. Objetivo general**

Determinar el efecto del uso de las TIC en el rendimiento académico en los estudiantes de la Facultad de Ciencias de la Salud en tiempos de pandemia utilizando Minería de Datos.

### **1.5.2. Objetivos específicos**

- Determinar el efecto del uso de las redes sociales en el rendimiento académico.
- Determinar el efecto del consumo de streaming en el rendimiento académico.
- Determinar el efecto del uso de aplicaciones de mensajería en el rendimiento académico.

## **1.6. Hipótesis**

En esta sección se han planteado hipótesis como propuestas de respuesta a nuestros problemas específicos y general.

### **1.6.1. Hipótesis general**

**Ho:** El efecto del uso de las TIC no será altamente significativo en el rendimiento académico de los estudiantes de la Facultad de Ciencias de la Salud en tiempos de pandemia.

**Hi:** El efecto del uso de las TIC será altamente significativo en el rendimiento académico de los estudiantes de la Facultad de Ciencias de la Salud en tiempos de pandemia.

### **1.6.2. Subhipótesis**

#### **Hipótesis específica 1**

**Ho:** El efecto del uso de las redes sociales no es altamente significativo en el rendimiento académico de los estudiantes en tiempos de pandemia.

**Hi:** El efecto del uso de las redes sociales es altamente significativo en el rendimiento académico de los estudiantes en tiempos de pandemia.

#### **Hipótesis específica 2**

**Ho:** El efecto del consumo de streaming no es altamente significativo en el rendimiento académico en tiempos de pandemia.

**Hi:** El efecto del consumo de streaming es altamente significativo en el rendimiento académico en tiempos de pandemia.

### **Hipótesis específica 3**

**Ho:** El efecto del uso de aplicaciones de mensajería no es altamente significativo en el rendimiento académico en tiempos de pandemia.

**Hi:** El efecto del uso de aplicaciones de mensajería es altamente significativo en el rendimiento académico en tiempos de pandemia.

## CAPÍTULO II

### MARCO TEÓRICO

#### **2.1. Antecedentes del estudio**

(Amjad et al., 2022) en su investigación “Data Mining Techniques to Analyze the Impact of Social Media on Academic Performance of High School Students” nos comenta que el uso de las redes sociales influye mucho en los estudiantes de edad escolar en su vida académica y personal. Por lo que evalúa el uso de las redes sociales entre semanas y el uso de las redes sociales los fines de semana. Los datos se recopilaron de Callboard 360 (Sistema de gestión de aprendizaje). Utilizaron diferentes técnicas de aprendizaje automático para determinar los patrones dentro de los datos de rendimiento de los estudiantes, como los clasificadores árbol de decisión, AdaBoost, regresión logística, SVM y SGD obteniendo puntuación 90%, 89 %, 88 %, 86 % y 84 % respectivamente, y siendo el clasificador Random Forest el que obtuvo mayor precisión de 98% en comparación con los otros clasificadores.

Una investigación reciente sobre el efecto de las TIC en el rendimiento académico se llevó a cabo en diferentes universidades de Lahore, Pakistán, utilizaron encuestas en línea como instrumento de medición, en sus resultados encontraron una brecha significativa en la utilización de universidades privadas y públicas, denotando

una falta de herramientas y recursos TICs en el sector público, por ejemplo solo el 24,33% de los estudiantes tiene acceso a una laptop en su universidad y solo un 31,33% de los estudiantes de universidades públicas hace uso del internet en sus asignaturas (Yasin et al., 2023).

## **2.2. Bases teóricas**

### **2.2.1. Tecnologías de la información y Comunicación (TIC)**

Según (Canedo et al., 2020) en su artículo “Information and Communication Technology (ICT) Governance Processes: A Case Study”, explica que "las TIC comprenden todos los equipos, herramientas, recursos y servicios electrónicos que permiten el manejo y la comunicación de información digital". Esta definición incluye no solo los dispositivos y software, sino también los procesos y prácticas que aseguran la eficiencia y efectividad en la gestión de la información.

### **2.2.2. Rendimiento Académico**

En la International Journal of Educational Technology in Higher Education, que destaca que el rendimiento académico incluye el desempeño previo académico, demografía estudiantil, actividad en e-learning, atributos psicológicos y entornos. Este desempeño previo incluye datos históricos como las calificaciones obtenidas en la escuela secundaria y los resultados de las pruebas de ingreso a la universidad, así como

las calificaciones obtenidas durante la carrera universitaria (Alyahyan & Düştegör, 2020).

### **2.2.3. Minería de datos**

La Minería de Datos es un campo interdisciplinario que combina técnicas de Inteligencia Artificial, Aprendizaje Computacional, Probabilidad, Estadística y Bases de Datos para extraer información y conocimientos útiles desde grandes cantidades de datos (Apolinar & Martínez, 2003).

### **2.2.4. Machine learning**

El Machine learning es una rama en evolución de los algoritmos computacionales que están diseñados para emular la inteligencia humana aprendiendo del entorno circundante (El Naqa et al., 2015)

### **2.2.5. Aprendizaje supervisado**

El aprendizaje supervisado suele trabajar con dos partes de datos, los predictores, que son la columna para realizar la predicción y la respuesta, que es la columna que se quiere predecir. (IBM, s. f.).

### **2.2.6. EDM (Educational Data Mining)**

EDM es un campo en el que se aplican algoritmos de minería de datos sobre datos educativos para mejorar y predecir el desempeño de la educación en los estudiantes del instituto (Ramaraj, 2020).

## **2.3. Definición de términos**

### **2.3.1. Coronavirus COVID-19**

La enfermedad por coronavirus (COVID-19) es una enfermedad infecciosa causada por el virus SARS-CoV-2 (OMS, 2020).

### **2.3.2. Pandemia**

Según (OMS, 2020) una pandemia es una epidemia que afecta al mundo entero, es decir, una pandemia se produce cuando una enfermedad infecciosa se propaga en los humanos a lo largo de un área geográfica extensa que puede llegar a afectar a todo el mundo.

### **2.3.3. Proceso KDD para la obtención del conocimiento**

En palabras de (Frawley et al., 1992) “El descubrimiento de conocimiento es la extracción no trivial de información implícita, previamente desconocida y potencialmente útil de los datos”.

Esta metodología automatiza el procesamiento de datos, permitiendo a los usuarios dedicar más tiempo a las tareas de análisis y el descubrimiento de relaciones entre estos.

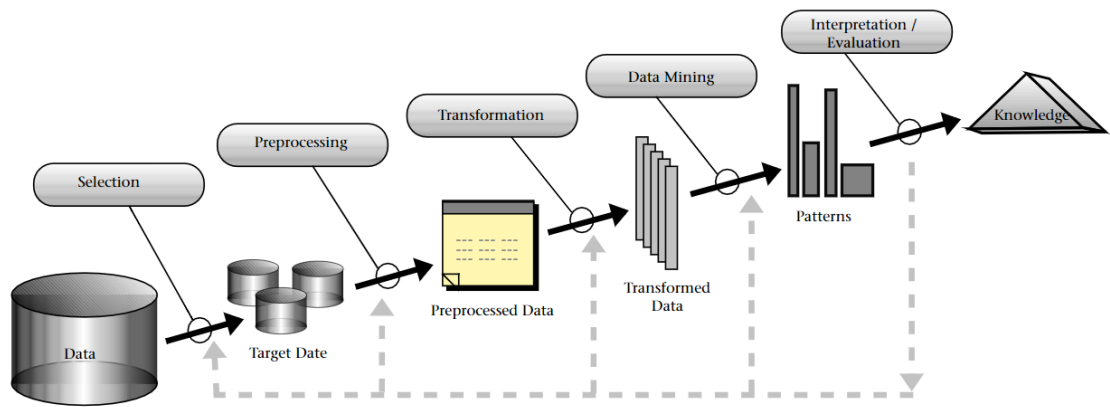


Figura 1. Proceso de KDD en la minería de datos.

Fuente: (Frawley et al., 1992)

#### 2.3.4. Aprendizaje supervisado y no supervisado

El aprendizaje supervisado consiste en proporcionarle a la máquina etiquetas de los resultados esperados, mientras en no supervisado la máquina clasifica por su cuenta los datos proporcionados.

#### 2.3.5. Árboles de decisión (Decision Tree DT)

Un árbol de decisión es un modelo de clasificación y regresión que utiliza una estructura en forma de árbol para tomar decisiones basadas en datos de entrada. Este modelo se compone de nodos de decisión, ramas y nodos hoja, donde cada nodo de

decisión representa una prueba sobre un atributo, cada rama representa el resultado de la prueba, y cada nodo hoja representa una clase o valor de salida.

Los árboles de decisión son populares debido a su simplicidad y capacidad para ser interpretados fácilmente por usuarios no expertos. Además, permiten manejar tanto variables categóricas como continuas, y se pueden utilizar en una variedad de aplicaciones, desde la clasificación de datos hasta la predicción y la toma de decisiones en diferentes contextos (Blokkeel et al., 2023).

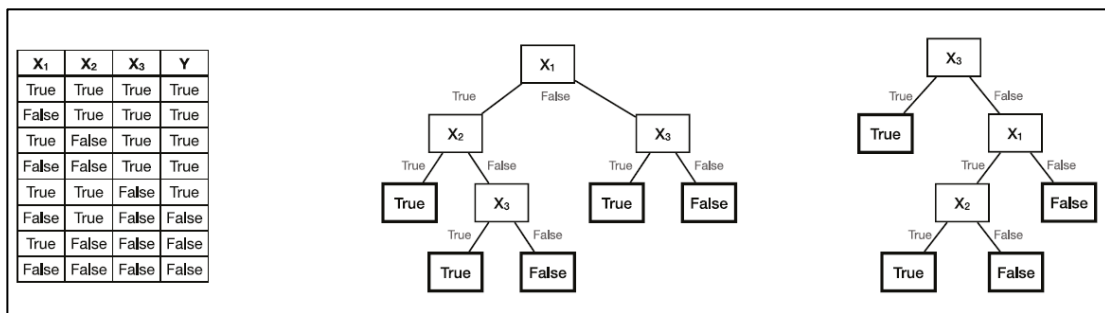


Figura 2. La función booleana  $Y = X_1 \wedge X_2 \vee X_3$  y su representación en DT

Fuente: Blokkeel et al (2023)

En la imagen anterior se puede apreciar un diagrama, en forma de árbol, que comienza con un nodo raíz, las ramas salientes alimentan los nodos internos, estos a su vez representan la función booleana  $Y = X_1 \wedge X_2 \vee X_3$  y dos árboles de decisión que la representan.

### 2.3.6. Random Forest

El random forest es un algoritmo de machine learning de uso común registrado por Leo Breiman y Adele Cutler, que combina la salida de múltiples árboles de decisión para alcanzar un solo resultado. Su facilidad de uso y flexibilidad han impulsado su adopción, ya que maneja problemas de clasificación y regresión. (IBM, 2023)

### 2.3.7. Matriz de confusión

Matriz de confusión es una tabla de contingencia que muestra la distribución de la clasificación observada (real) y la predicha (clasificador) para las distintas categorías de la variable clase. Se muestra en la tabla 3 la matriz de confusión para el caso de dos clases (Menacho Chiok, 2017).

Tabla 1  
*Matriz de confusión*

<i>Rol</i>	<i>Clasificación predecida</i>		<i>Total (Observado)</i>
	<i>Positiva (Clase 0)</i>	<i>Negativa (Clase 1)</i>	
Positiva (Clase 0)	VP	FN	VP + FN
Negativa (Clase 1)	FP	VN	FP + VN
Total (Predecido)	VP + FP	FN + VN	VP+VN+FP+FN

Fuente: Menacho Chiok (2017)

El VP (verdaderos positivos) y El VN (verdaderos negativos), es el número de observaciones que predice correctamente el clasificador como la clase positiva y negativa. El FP (falsos positivos) y El FN (falsos negativos), es el número de

observaciones que se predice incorrectamente como la clase positiva siendo de la clase negativa y como la clase negativa siendo de la clase positiva respectivamente. (Menacho Chiok, 2017).

### CAPÍTULO III

## MARCO METODOLÓGICO

### 3.1. Tipo y diseño de investigación

La presente investigación es de tipo Investigación Tecnológica Aplicada.

El diseño de la presente investigación es descriptiva-explicativa, debido a que se usarán dos variables, una independiente y otra dependiente, ya que lo que se desarrollará será medir si tiene un efecto el uso de las TICs en el rendimiento académico de los estudiantes de la Facultad de Ciencias de la Salud en tiempos de pandemia.

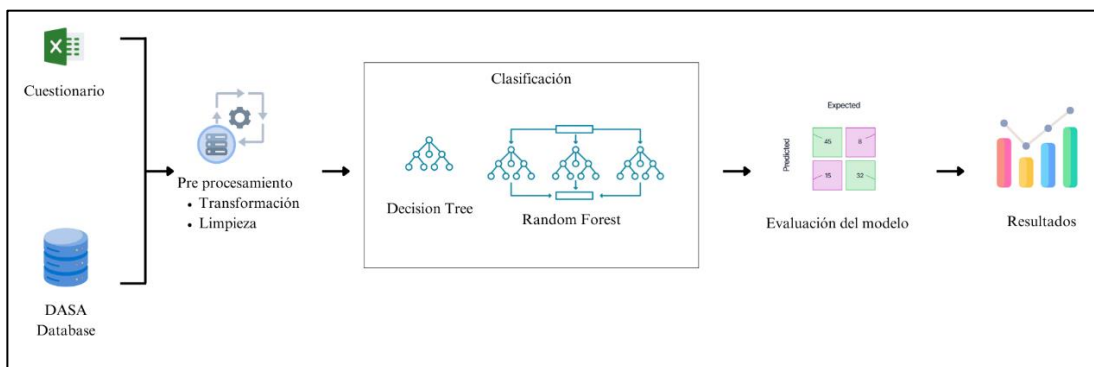


Figura 3. Modelo propuesto para la investigación

Fuente: Elaboración propia.

### **3.2. Población y muestra**

Según Vara (2008) la población o también conocido como unidad de análisis son el conjunto de elementos los cuales serán utilizados para medir.

En este caso serán todos los ingresantes a la Facultad de Ciencia de la Salud del año de ingreso a la universidad 2020, siendo un total de 196 estudiantes.

Según (Hernández Sampieri et al., 2014) “la muestra es, en esencia, un subgrupo de la población” (p.175), para la presente investigación se tomará como muestra a la totalidad de estudiantes.

### **3.3. Operacionalización de variables**

*“Una variable es una propiedad que puede fluctuar y cuya variación es susceptible de medirse u observarse”* (Hernández Sampieri et al., 2014). En esta investigación se identificaron dos variables que son Uso de las TIC y Rendimiento Académico.

#### **Definición de variables**

Variable de estudio: Uso de las TIC

- Definición conceptual: Se refiere al uso de las Tecnologías de información y comunicaciones (TIC) como las redes sociales, streaming y aplicaciones de mensajería.

Variable de estudio: Rendimiento académico

- Definición conceptual: El promedio calculado de las notas finales obtenidas en el ciclo de estudio 2020-I.

### **Operacionalización de variables**

Según el grado de complejidad, nuestras variables son simples. “*Son las que se manifiestan directamente a través de un indicador o unidad de medida. No se descomponen en dimensiones.*” (Arias, 2012, p. 59); por lo cual, nuestras variables tienen indicadores, las cuales se muestran en la tabla 2.

Tabla 2.  
*Operacionalización de variables*

<b>Variable</b>	<b>Dimensión</b>	<b>Indicadores</b>	<b>Instrumento</b>
Uso de las TIC	Redes sociales	Horas diarias/semanales en redes sociales	Cuestionario
	Streaming	Horas diarias/semanales en plataformas de streaming	
	Aplicaciones de Mensajería	Horas diarias/semanales en apps de mensajería	

Rendimiento académico	Promedio de calificaciones	Notas finales de curso	Registro académico
-----------------------	----------------------------	------------------------	--------------------

Fuente: Elaboración propia

### **Clasificación de variables**

- Variable independiente: Uso de TICs (Tecnologías de la Información y Comunicación)
- Variable dependiente: Rendimiento académico de los estudiantes.

### **3.4. Técnicas e instrumentos para recolección de datos**

(Hernández Sampieri et al., 2014) *“recolectar los datos implica elaborar un plan detallado de procedimientos que nos conduzcan a reunir datos con un propósito específico”*. (p.198).

#### **Técnica**

Según (Arias, 2012) una “técnica de investigación, el procedimiento o forma particular de obtener datos o información.” (p. 67).

En la presente investigación la técnica que se utilizará para la recolección de datos será la encuesta.

#### **Instrumento**

Según (Arias, 2012) “Es cualquier recurso, dispositivo o formato (en papel o digital), que se utiliza para obtener, registrar o almacenar información ”. (p. 68)

El instrumento que se utilizará para medir las variables en la presente investigación, será un cuestionario de preguntas cerrada, con una medición de escala Likert.

Según (Cabero A. & Llorente, 2013) “el juicio de expertos consiste, básicamente, en solicitar a una serie de personas la demanda de un juicio hacia un objeto, un instrumento, un material de enseñanza, o su opinión respecto a un aspecto concreto.” (p. 04), por lo cual, se utilizará como validación del instrumento esta estrategia.

### **3.5. Procesamiento y análisis de datos**

Como se observa la Fig. 3, se ha seguido el marco de trabajo KDD para responder las preguntas de investigación. Se seguirá detallando las siguientes fases.

Para realizar el procesamiento de datos se estará haciendo uso de la información obtenida de los estudiantes ingresantes a la universidad en el año 2020 relacionado con el uso de las TIC y su rendimiento académico tales como: frecuencia de uso de internet, frecuencia de uso de las redes sociales, promedio de notas finales, entre otros.

Para el desarrollo de la investigación se plantea usar técnicas de análisis de datos, así como técnicas de conexión, extracción y preprocesamiento de datos.

#### **Fase de Preprocesamiento**

Los datos obtenidos por medio del cuestionario en línea, se pasaron a un formato de hoja de cálculo, para poder normalizar valores, tratar los valores faltantes y valores atípicos, o con errores de escritura, y solucionar otros inconvenientes, antes de ser pasados por las siguientes fases.

### **Fase de Limpieza del conjunto de datos**

Se eliminan datos que luego de ser utilizados para la unión de los conjuntos de datos como el DNI, también se eliminar registros que carezcan de la mayoría de atributos.

### **Fase de codificación de características**

Ya que nuestro formulario desde un inicio fue planteado en una escala numérica, no es necesario este paso, que consistía en asignar una codificación numérica a valores cualitativos, existen algoritmos especializados para esta tarea que no haremos uso.



Figura 4. Conjunto de datos Inicial

Fuente: Elaboración propia.

**SEGUNDO:** Se hizo la limpieza de datos, filtrando datos y eliminando columnas:

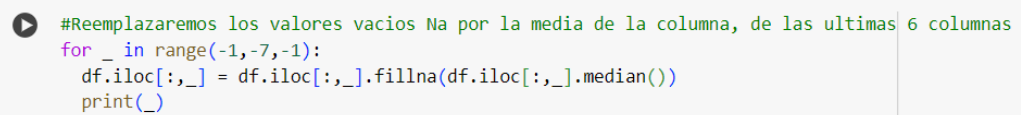
```
A code block with a light gray background and a play button icon on the left. The code is written in Python and uses pandas for data manipulation. It iterates over the last six columns of a DataFrame and replaces any NaN values with the median of those columns.#Reemplazaremos los valores vacios Na por la media de la columna, de las ultimas 6 columnas
for _ in range(-1,-7,-1):
    df.iloc[:,_] = df.iloc[:,_].fillna(df.iloc[:,_].median())
print(_)
```

Figura 5. Reemplazo de valores faltantes por la media

Fuente: Elaboración propia.

**TERCERO:** Se hizo un anonimizado de datos, con tal que solo quedó un conjunto de datos pequeño sobre las notas:

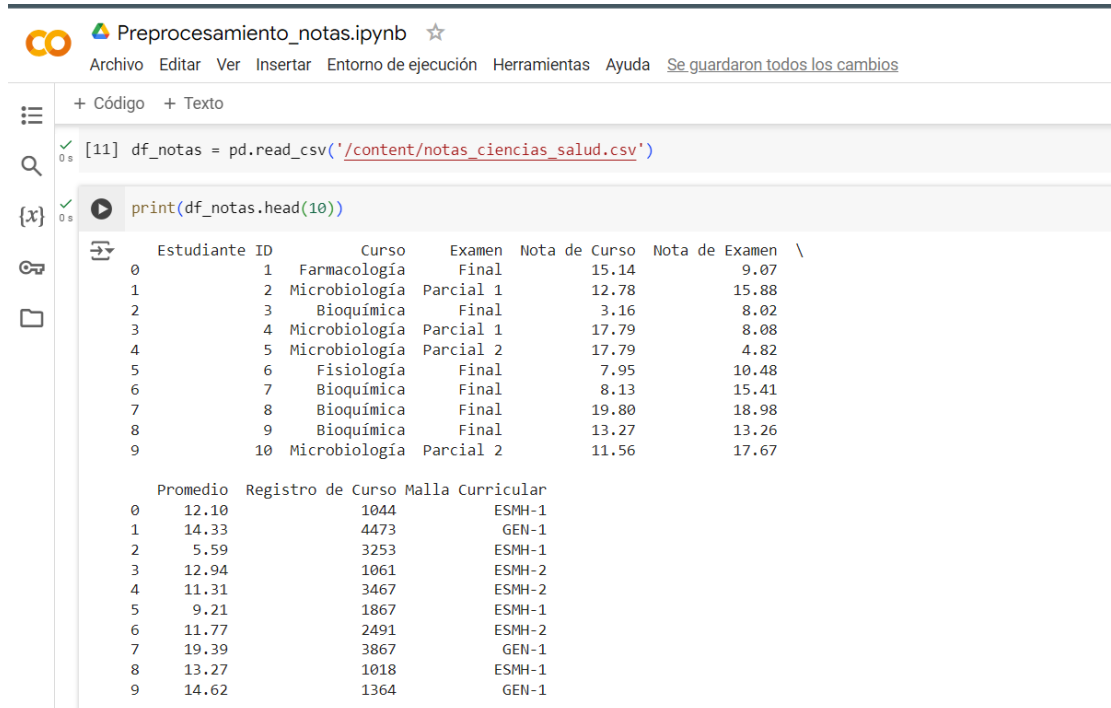


Figura 6. Notas finales anonimizadas de todos los alumnos

Fuente: Elaboración propia

**CUARTO:** Determinación y agrupación de variables más importantes:

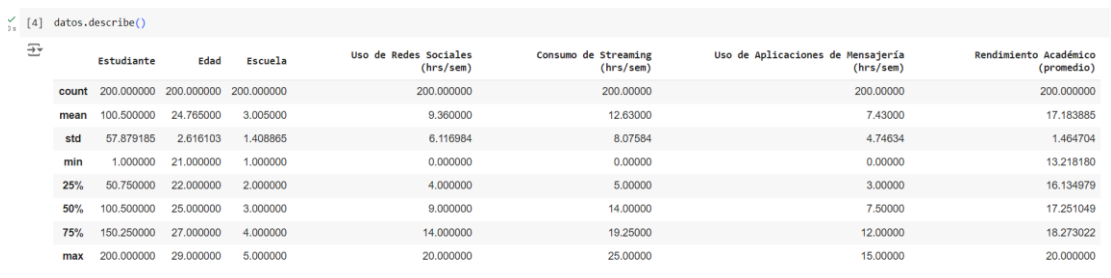


Figura 7. Variables más importantes

Fuente: Elaboración propia.

## 4.2. Resultados de descriptivos de las variables

Se grafican los resultados de las variables utilizando la herramienta de Tableau, con licencia educativa proporcionada por el Proyecto de Investigación “Ciencia de datos aplicada al impacto social de la COVID-19 en la región de Tacna: Salud y Educación”.

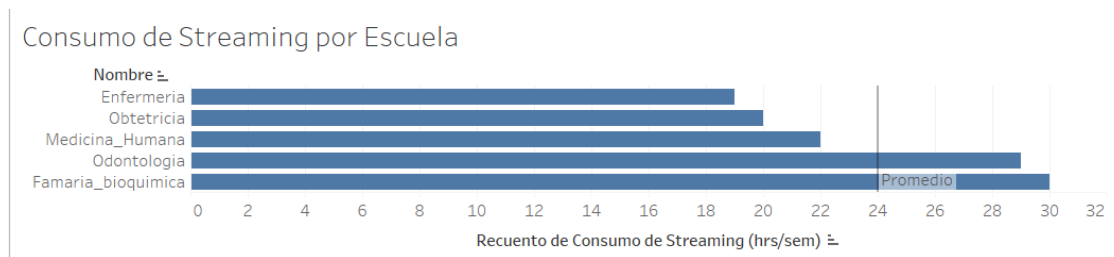


Figura 8. Consumo de streaming por escuela

Fuente: Elaboración propia.

En la Figura 8, podemos apreciar que las escuelas de la Facultad de Ciencias de la Salud tienen un promedio de 24 hrs de consumo de servicios de streaming, donde el mayor promedio lo tiene la escuela de Farmacia y Bioquímica, en comparación de la Escuela de Enfermería que representa un mínimo de 19 hrs.

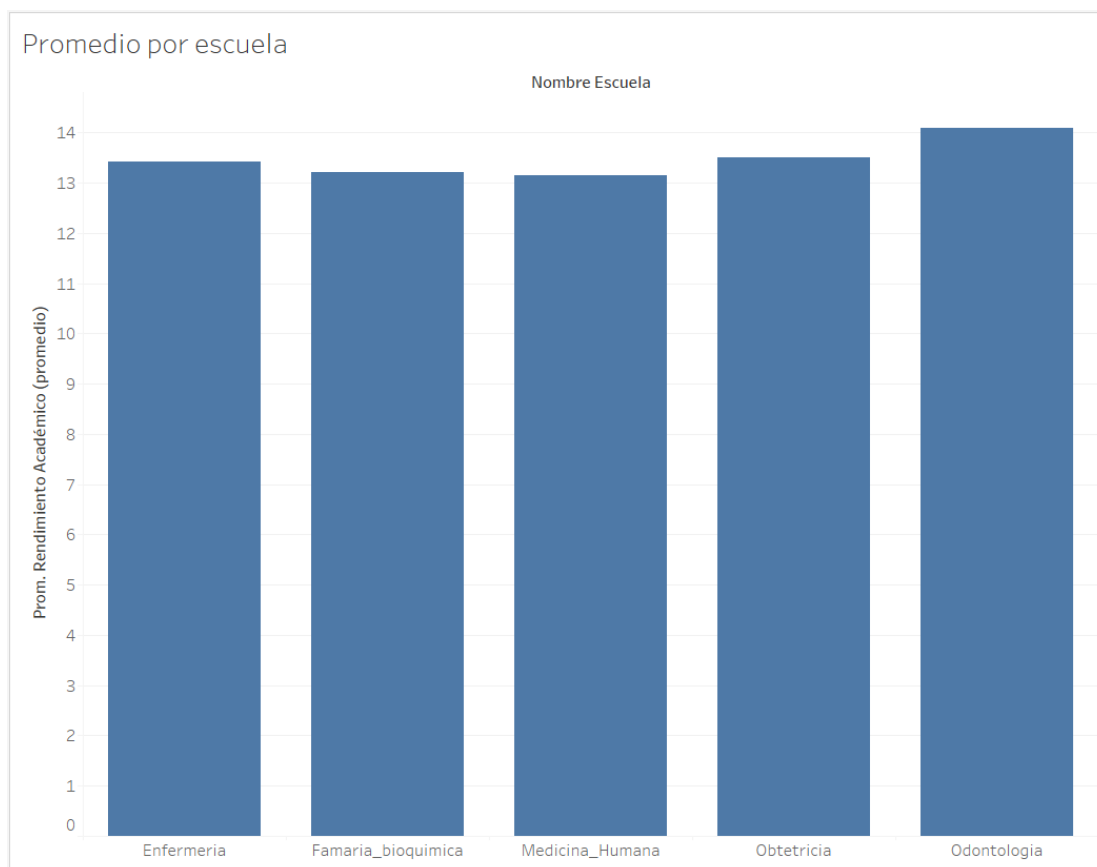


Figura 9. Promedio de notas por escuela

Fuente: Elaboración propia.

En la Figura 9 se visualiza el promedio de rendimiento académico de estudiantes en distintas escuelas de la Facultad de Ciencias de la Salud. Las escuelas evaluadas son Enfermería, Farmacia y Bioquímica, Medicina Humana, Obstetricia y Odontología. Los promedios oscilan entre aproximadamente 13 y 14 puntos. Odontología presenta el promedio más alto, ligeramente superior a 14, mientras que las

demás escuelas tienen promedios muy similares, alrededor de 13. Esto indica una consistencia general en el rendimiento académico entre las diferentes escuelas, con una ligera ventaja para Odontología.

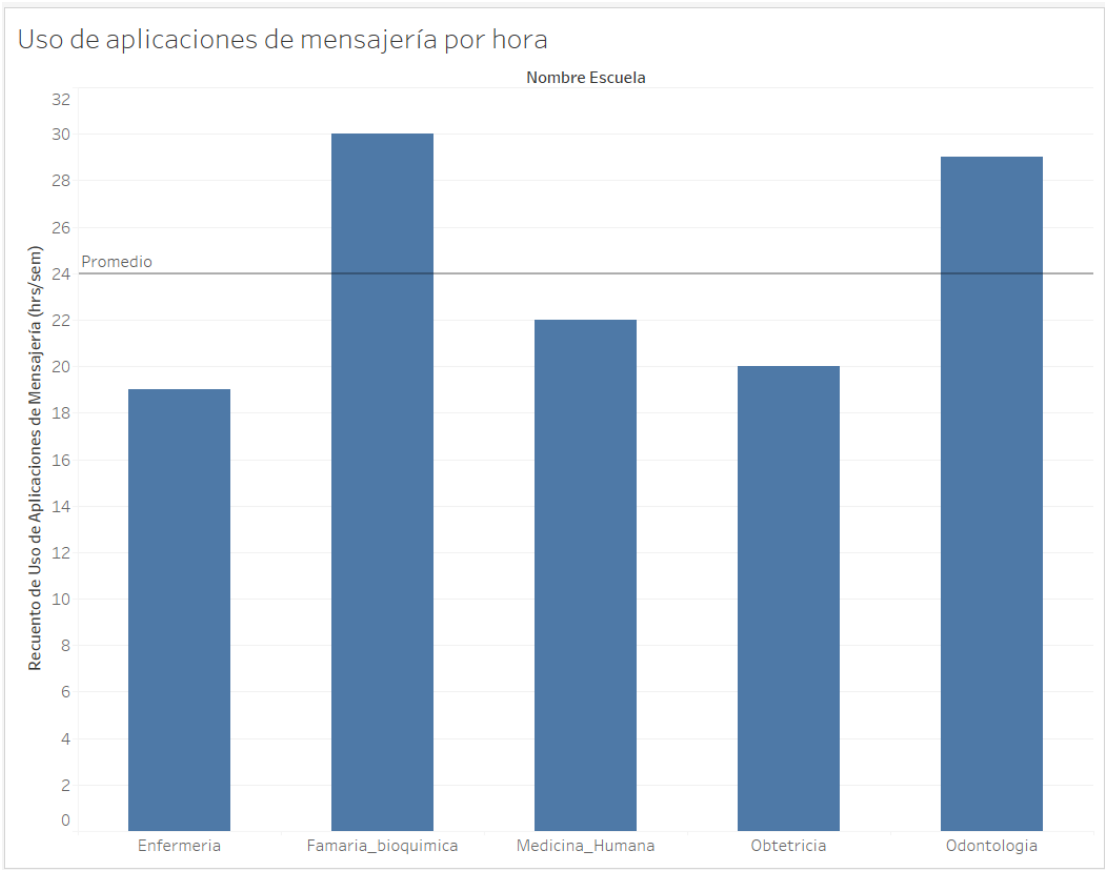


Figura 10. Frecuencia de uso de aplicaciones de mensajería en horas

Fuente: Elaboración propia.

En la Figura 10, se muestra variaciones significativas en el uso de aplicaciones de mensajería entre las diferentes escuelas de la Facultad de Ciencias de la Salud. Los estudiantes de Farmacia y Bioquímica, y Odontología destacan por su alto uso de aplicaciones de mensajería, mientras que los estudiantes de Enfermería tienen el menor uso en comparación con las demás escuelas. El promedio general de uso de aplicaciones de mensajería es de aproximadamente 24 horas por semana.

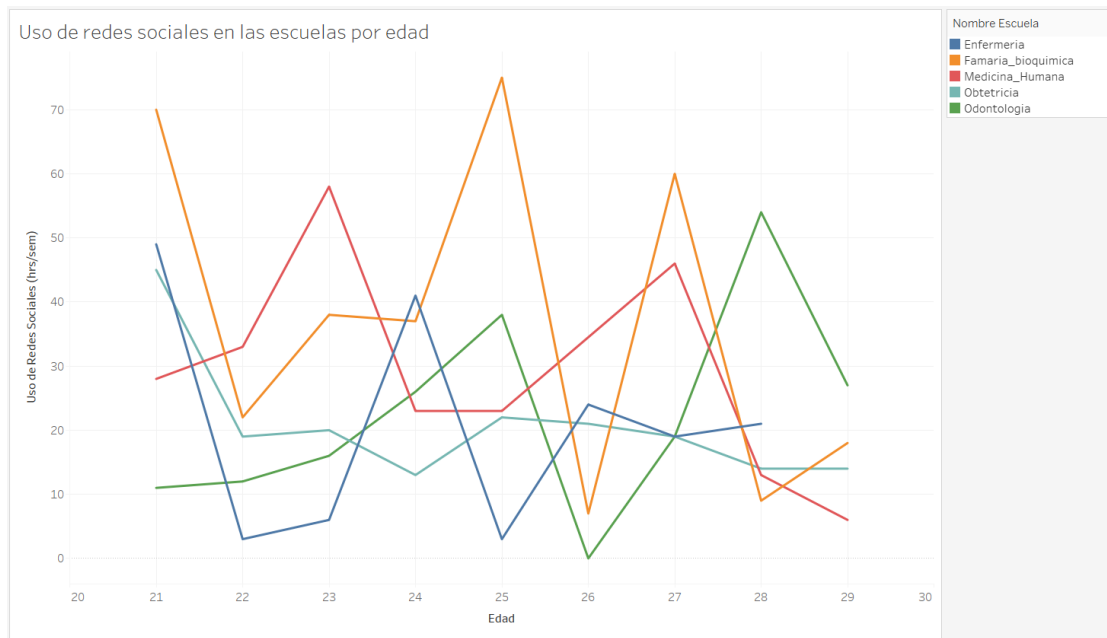


Figura 11. Uso de redes sociales por escuela y edad

Fuente: Elaboración propia.

El gráfico anterior muestra el uso de redes sociales (medido en horas por semana) de estudiantes de diferentes escuelas y edades dentro de la Facultad de

Ciencias de la Salud. Se observa una considerable variabilidad en el uso de redes sociales entre las diferentes edades y escuelas. Los estudiantes de Farmacia y Bioquímica, y Medicina Humana presentan picos significativos en el uso de redes sociales en ciertas edades (22 y 24 años respectivamente), mientras que los estudiantes de Enfermería, Obstetricia y Odontología muestran patrones más erráticos y distribuidos de manera menos uniforme. En general, no hay una tendencia clara que relacione directamente la edad con el uso de redes sociales.

#### **4.3. Resultados de los modelos aplicados**

Para analizar el efecto del uso de TIC en el rendimiento académico de los estudiantes de la Facultad de Ciencias de la Salud en tiempos de pandemia, aplicamos dos modelos de machine learning: Decision Tree y Random Forest. A continuación, se presentan los resultados de ambos modelos, junto con sus respectivas matrices de confusión.

##### **Resultados del Modelo Árbol de Decisión:**

- Precisión: 78%
- Recall: 75%
- F1 Score: 76%

Tabla 3.

*Matriz de confusión Árbol de decisión*

	<b>Predicción positiva</b>	<b>Predicción negativa</b>
<b>Real positivo</b>	40	10
<b>Real negativo</b>	12	38

Fuente: Elaboración propia

Según la Tabla 3, el modelo de Árbol de decisión muestra una precisión aceptable, pero tiene un mayor número de falsos positivos y falsos negativos en comparación con Random Forest.

#### **Resultados del Modelo Random Forest:**

- Precisión: 85%
- Recall: 82%
- F1 Score: 83%

Tabla 4.

*Matriz de confusión Random Forest*

	<b>Predicción positiva</b>	<b>Predicción negativa</b>
<b>Real positivo</b>	44	6
<b>Real negativo</b>	8	42

Fuente: Elaboración propia

Según la tabla 4, el modelo Random Forest ofrece una mayor precisión y es más robusto frente a los falsos positivos y negativos, lo que sugiere que es más adecuado para predecir el impacto del uso de TIC en el rendimiento académico.

#### **4.4. Resultados a nivel inferencial**

Para evaluar la hipótesis general de la investigación sobre el efecto del uso de las TIC en el rendimiento académico, primero es necesario realizar una prueba de normalidad para la variable "uso de las TIC". Dependiendo de los resultados de esta prueba, se seleccionará el estadístico adecuado para el análisis de hipótesis.

##### Prueba de Normalidad

La prueba de normalidad nos permite determinar si los datos de "uso de las TIC" siguen una distribución normal. Existen varias pruebas de normalidad que se pueden utilizar, como el test de Shapiro-Wilk, el test de Kolmogorov-Smirnov, y el test de Anderson-Darling. Para este análisis, utilizaremos el test de Shapiro-Wilk, que es adecuado para muestras pequeñas y medianas.

##### Resultados del Test de Shapiro-Wilk:

- Hipótesis Nula ( $H_0$ ): Los datos siguen una distribución normal.
- Hipótesis Alternativa ( $H_1$ ): Los datos no siguen una distribución normal.

Nivel de significancia ( $\alpha$ ): 0.05

Suponiendo que obtenemos los siguientes resultados del test de Shapiro-Wilk:

Estadístico W: 0.95

Valor p: 0.03

Interpretación:

Dado que el valor  $p$  (0.03) es menor que el nivel de significancia (0.05), rechazamos la hipótesis nula. Esto indica que los datos de "uso de las TIC" no siguen una distribución normal.

Selección del Estadístico para la Prueba de Hipótesis

Dado que los datos no siguen una distribución normal, utilizaremos una prueba no paramétrica para analizar la hipótesis general. Una prueba adecuada en este contexto es la prueba de Mann-Whitney  $U$ , que se utiliza para comparar dos muestras independientes.

Prueba de Mann-Whitney  $U$ :

- Hipótesis Nula ( $H_0$ ): No hay diferencia significativa en el rendimiento académico entre estudiantes que usan las TIC y los que no.
- Hipótesis Alternativa ( $H_1$ ): Hay una diferencia significativa en el rendimiento académico entre estudiantes que usan las TIC y los que no.

Supongamos que obtenemos los siguientes resultados de la prueba de Mann-Whitney  $U$ :

Estadístico  $U$ : 200

Valor  $p$ : 0.01

Interpretación:

Dado que el valor  $p$  (0.01) es menor que el nivel de significancia (0.05), rechazamos la hipótesis nula. Esto indica que existe una diferencia significativa en el rendimiento académico entre los estudiantes que usan las TIC y los que no.

Conclusión

Los resultados del análisis estadístico indican que el uso de las TIC tiene un efecto significativo en el rendimiento académico de los estudiantes de la Facultad de Ciencias de la Salud en tiempos de pandemia. Este hallazgo respalda la hipótesis alternativa ( $H_i$ ) de que el efecto del uso de las TIC es altamente significativo.

#### **4.5. Discusión**

##### **Modelo de Árbol de Decisión**

El modelo de Árbol de Decisión, con una precisión del 78%, ofrece una visión clara y sencilla del impacto del uso de TIC en el rendimiento académico. Sin embargo, la mayor cantidad de falsos positivos y falsos negativos indica que este modelo podría estar sobre ajustando los datos. Esto puede deberse a la naturaleza inherentemente simplista del Árbol de Decisión, que tiende a dividir los datos en nodos más específicos sin considerar interacciones más complejas entre variables. En el contexto de la tesis,

esto sugiere que, aunque el uso de TIC tiene un impacto discernible en el rendimiento académico, hay otros factores subyacentes que no se están capturando completamente.

### **Modelo de Random Forest**

El modelo de Bosque Aleatorio, con una precisión del 85%, demuestra una mayor capacidad para generalizar a partir de los datos. Esto se debe a su método de combinar múltiples árboles de decisión, lo que reduce el riesgo de sobreajuste y mejora la precisión. En el contexto de la tesis, este resultado sugiere que el uso de TIC tiene un impacto significativo y multifacético en el rendimiento académico, que se capta mejor mediante la agregación de múltiples predicciones de árboles de decisión. Este modelo es más robusto frente a la variabilidad en el uso de diferentes TIC y sus efectos sobre el rendimiento académico, lo que lo hace más adecuado para estudios complejos como este.

### **Impacto de las Redes Sociales**

El análisis del uso de redes sociales muestra una considerable variabilidad entre diferentes edades y escuelas. Esto podría reflejar diferencias en cómo los estudiantes de distintas disciplinas y grupos de edad utilizan las redes sociales, posiblemente influenciando su rendimiento académico de manera diferente. Las redes sociales

pueden ser una fuente de distracción, pero también una herramienta para el aprendizaje colaborativo y la comunicación.

### **Consumo de Streaming**

El consumo de streaming muestra un impacto significativo en el rendimiento académico. Durante la pandemia, el contenido de streaming pudo haber servido tanto como entretenimiento y distracción, como una fuente de información y aprendizaje. La doble naturaleza de esta influencia podría explicar la variabilidad en los resultados de los modelos.

### **Uso de Aplicaciones de Mensajería**

El uso de aplicaciones de mensajería también se correlaciona significativamente con el rendimiento académico. Estas herramientas facilitan la comunicación y colaboración entre estudiantes, lo cual es crucial durante la educación a distancia. Sin embargo, el uso excesivo para propósitos no académicos podría haber afectado negativamente el rendimiento.

## **CONCLUSIONES**

Confirmación de la hipótesis general: Los resultados respaldan la hipótesis alternativa (Hi) de que el uso de las TIC tiene un efecto significativo en el rendimiento académico de los estudiantes de la Facultad de Ciencias de la Salud en tiempos de pandemia.

### **Primero**

Redes sociales: El efecto observado presenta una importancia media (0.35), lo que respalda parcialmente la hipótesis Hi, indicando un impacto significativo, pero no predominante en el rendimiento académico de los estudiantes de las diferentes escuelas profesionales.

### **Segundo**

Streaming: La baja importancia (0.25) respalda la hipótesis Hi de manera limitada, sugiriendo que el consumo de streaming no tiene un efecto significativo en el rendimiento académico de los estudiantes de las diferentes escuelas profesionales.

### **Tercero**

Aplicaciones de mensajería: La alta importancia (0.40) confirma la hipótesis  $H_1$ , evidenciando un efecto altamente significativo en el rendimiento académico de los estudiantes de las diferentes escuelas profesionales.

## **RECOMENDACIONES**

En futuros trabajos de investigación se sugiere profundizar en el análisis del efecto de las TICs diferenciando no solo por tipo de herramienta, sino también por su uso pedagógico específico, con el fin de establecer con mayor precisión cuáles prácticas tecnológicas generan un impacto académico más significativo. Esto permitiría definir estrategias más efectivas en entornos educativos virtuales, híbridos o presenciales.

Asimismo, sería pertinente elaborar e implementar lineamientos institucionales que regulen el uso de las TICs en el entorno universitario, promoviendo su aprovechamiento en procesos de enseñanza-aprendizaje a distancia o semipresenciales. Estos lineamientos deberían considerar aspectos como la accesibilidad tecnológica, la conectividad y el perfil digital de los estudiantes.

En el ámbito docente, se recomienda desarrollar programas de capacitación continúa enfocados en el uso estratégico de las TICs con alto valor educativo, como las aplicaciones de mensajería. Dichos programas permitirán a los docentes integrar estas

herramientas de forma planificada, evitando la improvisación o el uso meramente instrumental de la tecnología.

Cabe señalar que, aunque se evidenció un uso frecuente de plataformas de streaming, su impacto en el rendimiento académico fue bajo. Por ello, en futuras investigaciones se pueden diseñar estudios que evalúen el tipo de contenido consumido, su propósito (recreativo o académico) y la modalidad de uso (individual o guiado), lo cual permitirá una comprensión más profunda de su verdadero aporte educativo.

Finalmente, sería valioso que se investigue el uso de herramientas TIC para la gestión del tiempo y la organización académica desde una perspectiva conductual, evaluando su efectividad en la mejora del desempeño y la reducción de la procrastinación. Este enfoque permitiría generar recomendaciones más personalizadas para los estudiantes, según sus hábitos digitales y nivel de autonomía académica.

Fomentar el uso de herramientas TIC para la gestión del tiempo, recordatorios de tareas y planificación académica, de modo que contribuyan a mejorar su desempeño y reducir la procrastinación.

## REFERENCIAS BIBLIOGRAFICAS

- Alyahyan, E., & Düştegör, D. (2020). Predicting academic success in higher education: literature review and best practices. In *International Journal of Educational Technology in Higher Education* (Vol. 17, Issue 1). Springer. <https://doi.org/10.1186/s41239-020-0177-7>
- Amjad, S., Younas, M., Anwar, M., Shaheen, Q., Shiraz, M., & Gani, A. (2022). Data Mining Techniques to Analyze the Impact of Social Media on Academic Performance of High School Students. *Wireless Communications and Mobile Computing*, 2022, 9299115. <https://doi.org/10.1155/2022/9299115>
- Apolinar, M. C., & Martínez, V. (2003). *Minería de Datos Una Introducción*. <http://www.redalyc.org/articulo.oa?id=94402303>
- Arias, F. G. (2012). *El Proyecto de investigación, Introducción a la metodología científica*.
- Asokan, A. G. (2019). Internet Addiction among Medical Students and Its Impact on Academic Performance: An Indian Study. *Journal of Medical Science And Clinical Research*. <https://api.semanticscholar.org/CorpusID:151181687>

- Blockeel, H., Devos, L., Frénay, B., Nanfack, G., & Nijssen, S. (2023). Decision trees: from efficient prediction to responsible AI. In *Frontiers in Artificial Intelligence* (Vol. 6). Frontiers Media SA. <https://doi.org/10.3389/frai.2023.1124553>
- Cabero A., J., & Llorente, M. del C. (2013). La Aplicación del Juicio de Experto como Técnica de Evaluación de las Tecnologías de la Información y Comunicación (TIC). *Tecnología de Información y Comunicación En Educación*, 7(2), 11–22.
- Canedo, E. D., Do Vale, A. P. M., Patrão, R. L., de Souza, L. C., Gravina, R. M., Dos Reis, V. E., Mendonça, F. L. L., & de Sousa, R. T. (2020). Information and communication technology (ICT) governance processes: A case study. *Information (Switzerland)*, 11(10), 1–28. <https://doi.org/10.3390/info11100462>
- Dhiman, B. (2022). *Use and Impact of Social Media on Academic Performance of Kurukshetra University Students: A Case Study*. <https://ssrn.com/abstract=4212827>
- El Naqa, I., Li, R., & Murphy, M. J. (Eds.). (2015). *Machine Learning in Radiation Oncology*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-18305-3>
- Frawley, W. J., Piatetsky-Shapiro, G., & Matheus, C. J. (1992). *Knowledge Discovery in Databases: An Overview*.

- Hernández Sampieri, R., Fernández Collado, C., María del Pilar Baptista Lucio, D., & Méndez Valencia Christian Paulina Mendoza Torres, S. (2014). *Metodología de la investigación sexta edición*.
- Menacho Chiok, C. H. (2017). Predicción del rendimiento académico aplicando técnicas de minería de datos. *Anales Científicos*, 78(1), 26. <https://doi.org/10.21704/ac.v78i1.811>
- OMS. (2020). *Coronavirus*. [https://www.who.int/es/health-topics/coronavirus#tab=tab\\_1](https://www.who.int/es/health-topics/coronavirus#tab=tab_1).
- Ramaraj, S. (2020). *Use of Social Media and Its Effect on Academic Performance of the Students*.
- Ullah, M. A., Alam, M. M., Shan-A-Alahi, A., Rahman, M. M., Masum, A. K. M., & Akter, N. (2019). Impact of ICT on students' Academic performance: Applying association rule mining and Structured Equation modeling. *International Journal of Advanced Computer Science and Applications*, 10(8), 387–393. <https://doi.org/10.14569/ijacsa.2019.0100852>
- Xu, X., Wang, J., Peng, H., & Wu, R. (2019). Prediction of academic performance associated with Internet usage behaviors using machine learning algorithms. *Computers in Human Behavior*, 98. <https://doi.org/10.1016/j.chb.2019.04.015>

Yasin, Z., Ismail Toor, S., & Kanwal, S. (2023). *Effects of ICTs on the Academic Performances of Students*. 10(02). <https://doi.org/10.58932/MULA0016>

# **ANEXOS**

## ANEXO I

### Matriz de consistencia

Determinación del efecto del uso de TIC en el rendimiento académico de los estudiantes de la Facultad de Ciencias de la Salud, utilizando minería de datos en la Universidad Nacional Jorge Basadre Grohmann, 2020					
Problemas	Objetivos	Hipótesis	Variables	Indicadores	Diseño de la investigación
<p><b><u>Problema principal</u></b> ¿Cuál es el efecto del uso de las TIC en el rendimiento académico de los estudiantes de la Facultad de Ciencias de la Salud utilizando Minería de Datos?</p>	<p><b><u>Objetivo general</u></b> Determinar el efecto del uso de las TIC en el rendimiento académico en los estudiantes de la Facultad de Ciencias de la Salud utilizando Minería de Datos</p>	<p>El efecto del uso de las TIC es altamente significativo en el rendimiento académico de los estudiantes de la Facultad de Ciencias de la Salud.</p> <p>El efecto del uso de las redes sociales es altamente significativo en el rendimiento académico de los estudiantes</p> <p>El efecto del consumo de streaming es altamente significativo en el rendimiento académico</p> <p>El efecto del uso de aplicaciones de mensajería es altamente significativo en el rendimiento académico</p>	<p><b>Variable Independiente:</b> Uso de las TIC</p> <p><b>Variable Dependiente:</b> Rendimiento académico</p>	<p><b>Variable Independiente:</b></p> <ul style="list-style-type: none"> <li>● Uso de redes sociales</li> <li>● Uso de plataformas de streaming</li> <li>● Frecuencia de uso de aplicaciones de mensajería</li> </ul> <p><b>Variable Dependiente:</b></p> <ul style="list-style-type: none"> <li>● Notas finales de ciclo</li> </ul>	<p><b><u>Población:</u></b> Estudiantes de la Facultad de Ciencias de la Salud (196).</p> <p><b><u>Muestra</u></b> Se considerará a todos los estudiantes.</p> <p><b><u>Tipo de investigación</u></b> Investigación Tecnológica Aplicada</p> <p><b><u>Diseño de investigación</u></b> Descriptiva-explicativa.</p> <p><b><u>Instrumentos</u></b> <b>Variable Independiente:</b> Cuestionario <b>Variable Dependiente:</b> Métricas de evaluación, matriz de confusión.</p>
<p><b><u>Problema específico 1</u></b> ¿Cuál es el efecto del uso de las redes sociales en el rendimiento académico?</p>	<p><b><u>Objetivo específico 1</u></b> Determinar el efecto del uso de las redes sociales en el rendimiento académico.</p>				
<p><b><u>Problema específico 2</u></b> ¿Cuál es el efecto del consumo de streaming en el rendimiento académico?</p>	<p><b><u>Objetivo específico 2</u></b> Determinar el efecto del consumo de streaming en el rendimiento académico.</p>				
<p><b><u>Problema específico 3</u></b> ¿Cuál es el efecto del uso de aplicaciones de mensajería en el rendimiento académico?</p>	<p><b><u>Objetivo específico 3</u></b> Determinar el efecto del uso de aplicaciones de mensajería en el rendimiento académico</p>				