

**UNIVERSIDAD NACIONAL DANIEL ALCIDES CARRIÓN**

**FACULTAD DE INGENIERÍA**

**ESCUELA DE FORMACIÓN PROFESIONAL DE INGENIERÍA DE SISTEMAS**

**Y COMPUTACIÓN**



**T E S I S**

**Modelo predictivo basado en minería de datos para la mejora en toma de  
decisiones en el consultorio de medicina rehabilitación en el Hospital  
Daniel Alcides Carrión Pasco, 2023**

**Para optar el título profesional de:**

**Ingeniero de Sistemas y Computación**

**Autores:**

**Bach. Deisy Jakelin CELIS TRINIDAD**

**Bach. Elena Lucero JAPA OSCATEGUI**

**Asesor:**

**Mg. Lisbeth Gisela NEGRETE CARHUARICRA**

**Cerro de Pasco – Perú - 2024**

**UNIVERSIDAD NACIONAL DANIEL ALCIDES CARRIÓN**

**FACULTAD DE INGENIERÍA**

**ESCUELA DE FORMACIÓN PROFESIONAL DE INGENIERÍA DE SISTEMAS**

**Y COMPUTACIÓN**



**T E S I S**

**Modelo predictivo basado en minería de datos para la mejora en toma de  
decisiones en el consultorio de medicina rehabilitación en el Hospital  
Daniel Alcides Carrión Pasco, 2023**

**Sustentada y aprobada ante los miembros del jurado:**

---

Mg. Melquiades Arturo TRINIDAD MALPARTIDA  
**PRESIDENTE**

---

Mg. Oscar Cleverio CAMPOS SALVATIERRA  
**MIEMBRO**

---

Mg. Pit Frank ALANIA RICALDI  
**MIEMBRO**

## INFORME DE ORIGINALIDAD Y SIMILITUD



Universidad Nacional Daniel Alcides Carrión  
Facultad de Ingeniería  
Unidad de Investigación

### INFORME DE ORIGINALIDAD N° 036-2024-UNDAC/UIFI

La Unidad de Investigación de la Facultad de Ingeniería de la Universidad Nacional Daniel Alcides Carrión en mérito al artículo 23° del Reglamento General de Grados Académicos y Títulos Profesionales aprobado en Consejo Universitario del 21 de abril del 2022, La Tesis ha sido evaluado por el software antiplagio Turnitin Similarity, que a continuación se detalla:

Tesis:

**MODELO PREDICTIVO BASADO EN MINERÍA DE DATOS PARA LA  
MEJORA EN TOMA DE DECISIONES EN EL CONSULTORIO DE MEDICINA  
REHABILITACIÓN EN EL HOSPITAL DANIEL ALCIDES CARRIÓN PASCO,  
2023**

Apellidos y nombres de los tesistas:

**Bach. CELIS TRINIDAD, Deisy Jakelin**

**Bach. JAPA OSCATEGUI, Elena Lucero**

Apellidos y nombres del Asesor:

**Mg. NEGRETE CARHUARICRA, Lisbeth Gisela**

Escuela de Formación Profesional

**Ingeniería Sistemas y Computación**

Índice de Similitud

7%

**APROBADO**

Se informa al decanato para los fines pertinentes:

Cerro de Pasco, 6 de febrero del 2024

  
UNIVERSIDAD NACIONAL DANIEL ALCIDES CARRIÓN  
FACULTAD DE INGENIERÍA  
UNIDAD DE INVESTIGACIÓN  
**Luis Villa Requis Carbajal**  
DOCTOR EN CIENCIAS - DIRECTOR

## **DEDICATORIA.**

A Dios, por permitirme conocer a personas con un corazón noble y culminar con éxito esta carrera.

A mis Padres, piedra angular de mi vida e impulso para seguir adelante.

A Walter, por su sabiduría, paciencia y por todo su apoyo incondicional.

A mi amiga la Sra. Raída, por salvarme los almuerzos y por todo su apoyo moral.

A mí misma por creer en mí, trabajando y estudiando a la vez, con resiliencia en las adversidades y con determinación, cumplo una meta más en mi vida.

### **Con gratitud infinita Deisy.**

A mis queridos padres, cuyo amor incondicional, sacrificio y constante aliento han sido la piedra angular de mi camino hacia la realización de este logro.

A mis queridos hermanos, han sido mi fuerza en los momentos más desafiantes de este viaje.

A Carlos, tu presencia reconfortante y tu apoyo emocional han sido mi ancla en los momentos de duda y fatiga durante esta travesía.

### **Con amor y gratitud eterna, Elena.**

## **AGRADECIMIENTO**

A los instructores de nuestra casa de estudios, ellos nos brindaron los conocimientos en cada semestre, gracias a ello, hoy en día podemos llevar a cabo sus enseñanzas en el ámbito laboral.

A nuestra asesora académico, la Ing. Lisbeth Negrete, cuya guía experta, paciencia y sabios consejos, fueron fundamentales para el desarrollo de este trabajo de investigación.

Al personal del hospital Daniel Alcides Carrión que brindaron su valiosa asistencia durante la recopilación de datos y la investigación.

Finalmente, nos gustaría expresar gratitud al señor Rubén Japa, cuya diligencia ha contribuido significativamente en la obtención de nuestro título universitario.

## RESUMEN

El trabajo de investigación que realice se titula: “Modelo Predictivo basado en minería de datos para la mejora en toma de decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023” El objetivo principal es desarrollar un modelo predictivo basado en la exploración de datos para mejorar la toma de decisiones en la clínica de medicina de rehabilitación del Hospital Daniel Alcides Carrión Pasco, 2023. El método utilizado es el diseño de investigación utilizado antes del experimento. Esta población incluye todos los datos de los pacientes atendidos durante el año 2022 en el Hospital Daniel Alcides Carrión Pasco, y se seleccionó aleatoriamente una muestra de 137 pacientes. En este proyecto de minería de datos, el modelo de minería de datos mejoró la toma de decisiones en un 92% y alcanzó un 76-100%, lo cual es excelente. En resumen, la clasificación de las variables de los pacientes en la Oficina de Medicina de Rehabilitación se determinó utilizando la tecnología de árbol de decisión J48, siendo la variable más influyente el diagnóstico informado por los financiadores para lograr una precisión del 100 %.

**Palabras Clave.** Modelo predictivo, Toma de decisiones.

## ABSTRACT

The research work I carried out is titled: "Predictive Model based on data mining to improve decision making in the Rehabilitation Medicine Clinic at the Daniel Alcides Carrión Pasco Hospital, 2023" The main objective is to develop a predictive model based on the exploration of data to improve decision making in the rehabilitation medicine clinic of the Daniel Alcides Carrión Hospital Pasco, 2023. The method used is the research design used before the experiment. This population includes all the data of the patients treated during the year 2022 at the Daniel Alcides Carrión Pasco Hospital, and a sample of 137 patients was randomly selected. In this data mining project, the data mining model improved decision making by 92% and reached 76-100%, which is excellent. In summary, classification of patient variables in the Office of Rehabilitation Medicine was determined using J48 decision tree technology, with the most influential variable being the diagnosis reported by funders to achieve 100% accuracy.

**Keywords.** Predictive model, Decision making.

## INTRODUCCIÓN

La Clínica de Medicina de Rehabilitación del Hospital Daniel Alcides Carrión Pasco ha podido utilizar el método CRISP DM para desarrollar un modelo predictivo que mejora la toma de decisiones, como se demuestra en este estudio mediante regresión lineal. El hospital cuenta con diversos departamentos que atienden a diferentes especialidades, siendo las más importantes medicina, pediatría, cirugía, urgencias, enfermería, ginecología, trastornos obsesivos, patología/estadística o rehabilitación. Cada departamento tiene procedimientos de gestión internos, y los departamentos de estadística y tecnología de la información ayudan con la gestión interna, el control del desempeño y la gestión de decisiones.

De las investigaciones en curso, el objetivo es procesar, analizar y predecir datos para la mejora de toma de decisiones utilizando la minería de datos, y con este fin la Oficina de Medicina de Rehabilitación llevará a cabo investigaciones.

Este estudio consta de los siguientes capítulos:

El Capítulo 1 "Comprende las siguientes partes: ficha y concreción del problema, determinación del examen, formulación del problema, formulación de objetivos, razones del examen y limitaciones del examen".

El Capítulo 2 " Incluye las siguientes partes: esencia estudiado teórico, observación de conceptos, filiación de supuestos y variables, y recientemente observación operativa".

El Capítulo 3 "cubre: tipos de investigación, métodos de investigación, diseño de investigación, poblaciones y muestras, métodos e instrumentos de recopilación de datos, métodos de análisis y procesamiento de datos, procesamiento de datos estadísticos, selección de herramientas de investigación, validación y confiabilidad, y pautas éticas".

Los cuatro capítulos "incluyen: Resultados y discusión que abarcan las siguientes secciones: leyenda del trabajo, disección e hermenéutica de resultados,



testimonio de hipótesis y jaleo de resultados". Finalmente, "Ver Conclusiones, Recomendaciones, Referencias y anexos"

**Los autores**

## ÍNDICE

DEDICATORIA.	
AGRADECIMIENTO	
RESUMEN	
ABSTRACT	
INTRODUCCIÓN	
ÍNDICE	

### CAPITULO I

#### PROBLEMA DE INVESTIGACIÓN

1.1. IDENTIFICACIÓN Y DETERMINACIÓN DEL PROBLEMA .....	1
1.2. DELIMITACIÓN DE LA INVESTIGACIÓN.....	2
1.3. FORMULACIÓN DEL PROBLEMA .....	3
1.3.1. Problema general .....	3
1.3.2. Problemas específicos .....	3
1.4. FORMULACIÓN DE OBJETIVOS.....	3
1.4.1. Objetivo general .....	3
1.4.2. Objetivos específicos.....	3
1.5. JUSTIFICACIÓN DE LA INVESTIGACIÓN.....	3
1.6. LIMITACIONES DE LA INVESTIGACIÓN .....	4

### CAPITULO II

#### MARCO TEÓRICO

2.1. ANTECEDENTES DE ESTUDIO .....	5
2.1.1. A nivel Internacional .....	5
2.1.2. A nivel Nacional.....	6
2.1.3. A nivel Local.....	9
2.2. BASES TEÓRICAS – CIENTÍFICAS .....	9
2.2.1. Minería de datos.....	9
2.2.2. Modelos de Minería de Datos .....	11
2.2.3. Modelo Predictivo .....	11
2.2.4. Métodos del Modelo Predictivo.....	13
2.2.5. Técnicas del Modelo Predictivo. ....	14

2.3. DEFINICIÓN DE TÉRMINOS BÁSICOS.....	15
2.4. FORMULACIÓN DE HIPÓTESIS .....	16
2.4.1. Hipótesis general.....	16
2.4.2. Hipótesis Específicas .....	16
2.5. IDENTIFICACIÓN DE VARIABLES.....	16
2.6. DEFINICIÓN OPERACIONAL DE VARIABLES E INDICADORES.....	17

### **CAPITULO III**

#### **METODOLOGÍA Y TÉCNICAS DE INVESTIGACIÓN**

3.1. TIPO DE INVESTIGACIÓN.....	18
3.2. NIVEL DE INVESTIGACIÓN .....	18
3.3. MÉTODOS DE INVESTIGACIÓN .....	18
3.4. DISEÑO DE INVESTIGACIÓN.....	19
3.5. POBLACIÓN Y MUESTRA .....	19
3.6. TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS .....	20
3.7. SELECCIÓN, VALIDACIÓN Y CONFIABILIDAD DE LOS INSTRUMENTOS DE INVESTIGACIÓN .....	20
3.8. TÉCNICAS DE PROCESAMIENTO Y ANÁLISIS DE DATOS. ....	20
3.9. TRATAMIENTO ESTADÍSTICO.....	20
3.10. ORIENTACIÓN ÉTICA FILOSÓFICA Y EPISTÉMICA.....	20

### **CAPITULO IV**

#### **RESULTADOS Y DISCUSIÓN**

4.1. DESCRIPCIÓN DEL TRABAJO DE CAMPO .....	21
4.2. PRESENTACIÓN, ANÁLISIS E INTERPRETACIÓN DE RESULTADOS .....	22
4.3. PRUEBA DE HIPÓTESIS.....	32
4.4. DISCUSIÓN DE RESULTADOS.....	34

CONCLUSIONES

RECOMENDACIONES

REFERENCIAS BIBLIOGRÁFICAS

ANEXOS

## ÍNDICE DE TABLAS

TABLA 1. DEFINICIÓN OPERACIONAL DE VARIABLES .....	17
TABLA 2. DIAGNÓSTICO PARA REHABILITACIÓN.....	22
TABLA 3. DICCIONARIO DE DATOS.....	23
TABLA 4. RESULTADO DE LAS TÉCNICAS DE ESTUDIO.....	32
TABLA 5. TABLA DE PREDICCIÓN.....	33

## ÍNDICE DE FIGURAS

FIGURA 1. DEFINICIÓN DE MINERÍA DE DATOS. ....	10
FIGURA 2.FASES DE LA MINERÍA DE DATOS.....	11
FIGURA 3.ESQUEMA GENERAL DEL MODELO PREDICTIVO .....	12
FIGURA 4.REPRESENTACIÓN DEL MODELO PREDICTIVO.....	13
FIGURA 5.UBICACIÓN.....	22
FIGURA 6.DATOS DE PACIENTES EN EL CONSULTORIO DE REHABILITACIÓN .24	
FIGURA 7.DATOS DE PACIENTES EN EL CONSULTORIO DE REHABILITACIÓN EN SPSS .....	24
FIGURA 8.ANÁLISIS DESCRIPTIVO DE PACIENTES EN EL CONSULTORIO DE REHABILITACIÓN EN SPSS.....	25
FIGURA 9.GRÁFICO DE BARRAS ANÁLISIS DESCRIPTIVO DE PACIENTES EN EL CONSULTORIO DE REHABILITACIÓN EN SPSS. ....	25
FIGURA 10.RED MLP.....	26
FIGURA 11.REPORTE DE MÉTRICAS. ....	27
FIGURA 12.REPORTE DE NAIVE BAYES .....	29
FIGURA 13.REPORTE DE ÁRBOL DE DECISIÓN .....	30
FIGURA 14.ÁRBOL DE DECISIÓN .....	31

## **CAPITULO I**

### **PROBLEMA DE INVESTIGACIÓN**

#### **1.1. Identificación y determinación del problema**

Los altos niveles competitivos que deben alcanzarse en las organizaciones de todos los sectores, obligan a sus responsables a apoyarse para la toma de decisiones en información directa, real, veraz y demostrable que evidencie la gestión realizada por cada uno de los miembros de la institución (Ministerio de Salud, 2013)

Los indicadores de gestión en el sector salud, son herramientas que permiten medir la administración que buscan optimizar la oferta de medicina rehabilitación a una demanda de necesidades de atención de salud que les permita tomar decisiones para la gestión administrativa eficiente y oportuna (Ministerio de Salud, 2013), bajo este esquema el Hospital Daniel Alcides Carrión Pasco que recibe una gran demanda de pacientes en el 2022 y 2023 se registraron en promedio 26137 casos atendidos.

Para la adecuada gestión de la Medicina de Rehabilitación, tenga un procesamiento de datos, controle sus indicadores y utilice estos para ayudar a tomar decisiones a nivel.

A modo de ejemplo, nuestro departamento de Medicina de Rehabilitación está teniendo serios problemas para procesar, analizar y pronosticar la información necesaria para tomar decisiones debido al reporte mensual de dichas métricas.

- Número de pacientes.
- Diagnóstico de enfermedades.
- Terapias.

De la información requerida, el departamento de Medicina de Rehabilitación (Número de pacientes, Diagnóstico de enfermedades, Terapias)

Si esta situación continúa, los centros de rehabilitación seguirán registrándose en forma de "informes de rehabilitación" con datos ilegibles, lo que dificulta el análisis y la previsión y no cooperará con la toma de decisiones.

## **1.2. Delimitación de la investigación**

### **a) Temporal**

En el proceso de investigación para la recolección de datos, el estudio de la interpretación de la información realizados con los datos del año 2022 y 2023.

### **b) Espacial**

Realizaremos nuestro trabajo de investigación utilizando los datos de pacientes de medicina rehabilitadora del Hospital Daniel Alcides Carrión Pasco, que se encuentran en una base de datos obtenida con un programa Excel.

### **c) Conceptual**

La investigación teórica se basa en modelos predictivos desarrollados examinando datos y mejorando las decisiones.

### **1.3. Formulación del problema**

#### **1.3.1. Problema general**

¿En qué medida el Modelo Predictivo basado en Minería de Datos mejora la toma de decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023?

#### **1.3.2. Problemas específicos**

- a) ¿Se podrá desarrollar el modelo predictivo para la mejora en toma de decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023?
- b) ¿Se podrá evaluar el modelo predictivo para la mejora en toma de decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023?

### **1.4. Formulación de objetivos**

#### **1.4.1. Objetivo general**

Desarrollar un modelo predictivo basado en minería de datos para la mejora de toma de decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023.

#### **1.4.2. Objetivos específicos**

- a) Desarrollar el modelo predictivo para la mejora en toma de decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023.
- b) Evaluar el modelo predictivo para la mejora en toma de decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023.

### **1.5. Justificación de la investigación**

El personal directivo del hospital que gestiona la toma de decisiones no tiene la información adecuada para gestionar los recursos internos del hospital, con la investigación se les brinda datos del área de medicina rehabilitación para



optimizar la administración de los servicios a nivel operativo, utilizando como medio la minería de datos la cual se fundamenta en la busca de patrones dentro de grandes bases de datos utilizando métodos predictivos, haciendo uso de recursos informáticos y tecnológicos.

El desarrollo del modelo predictivo basado en minería de datos permite analizar, procesar y predecir los indicadores que sirven para la toma de decisiones ya que con ello se tiene el acceso inmediato a los datos.

#### **1.6. Limitaciones de la investigación**

Con base en las características del tema y del estudio propuesto, se identificaron tres categorías de limitaciones:

- **Limitación temporal:** Los datos se recopilaron entre 2022 y 2023 para el estudio.
- **Limitación de información:** Las limitaciones encontradas para la realización de esta investigación se centran en el recelo de los responsables en compartir los datos y base de datos de los pacientes diagnosticados con rehabilitación, ya que como comprenderán es información confidencial.
- **Limitación económica:** El estudio realizado es autofinanciado, existiendo por consiguiente limitaciones en el gasto de los recursos necesarios para lograr los objetivos de la investigación.

## **CAPITULO II**

### **MARCO TEÓRICO**

#### **2.1. Antecedentes de estudio**

##### **2.1.1. A nivel Internacional**

Según Uvidia, Cisneros, Méndez y Villa (2019) en su trabajo de investigación titulado “Minería de datos para la toma de decisiones en la unidad de nivelación y admisión universitaria ecuatoriana” La minería de datos, también conocida como minería de datos (DM), utiliza la selección de algoritmos y el análisis de información para obtener patrones que, después de ser observados y probados, se convierten en conocimiento que ayuda a los departamentos de evaluación y admisiones de la educación superior a tomar decisiones. El ESPOCH es el Instituto Tecnológico de Chimborazo. Desde 2012, ESPOCH ha generado datos siguiendo los estándares actuales, lo que ha producido una gran cantidad de información sin valor para la toma de decisiones. Los algoritmos de predicción de clasificación (árboles de decisión y redes neuronales) y regresión (regresión lineal y optimización) se pueden analizar utilizando la minería de

datos del entorno de análisis de conocimiento de Waikato (WEKA). El objetivo es conocer las tendencias de uso por región, género y año hasta realizar predicciones. Finalmente, el análisis de parámetros estadísticos determinó los mejores algoritmos de confiabilidad de la información y generación de conocimiento para la toma de decisiones académicas, los cuales son: redes bayesianas y optimización de secuencia mínima. De un total de 12.529 entradas, el algoritmo de red bayesiana tuvo el mayor número de casos correctamente clasificados, con un total de 8.194, lo que representa el 65,4003%. La índice kappa del algoritmo de red bayesiano es el más cercano a 1, que es 0,3118, que es el valor con el mayor grado de superposición de variables, que es un punto aceptable. La red bayesiana tiene un error absoluto de 0,4402; cuanto menor sea, mejor porque la diferencia con la media exacta es menor.

Según García (2013) en su trabajo titulado “Aplicación de técnicas de Minería de Datos a datos obtenidos por el Centro Andaluz de Medio Ambiente (CEAMA)” Hoy día nuestra sociedad genera grandes cantidades de información que unido al aumento de las capacidades de almacenamiento, han hecho que todo tipo de organizaciones puedan disponer de una gran cantidad y variedad de datos relativos a su actividad diaria. Esta información es esencial para la toma de decisiones porque le da a la empresa una perspectiva (qué se está haciendo y cómo se está haciendo) y una visión de futuro (cómo se desarrollará la organización en el corto y medio plazo): La etapa de producción. Sin embargo, la mayoría de la información recopilada en la base de datos está mal estructurada y es difícil de usar desde un punto de vista estadístico, por lo que su uso requiere minería de datos, un proceso de procesamiento y análisis detallados de los datos recopilados.

### **2.1.2. A nivel Nacional**

Según Becerra y Villareal (2021) en su trabajo de investigación titulado “Data Mining para modelo predictivo de ventas y servicios de mantenimiento en

un concesionario Automotriz Ligero”. La competencia entre empresas en el segmento de vehículos de pasajeros suele ser feroz hoy en día, ya que los competidores desarrollan diferentes estrategias. Nuestra investigación tiene como objetivo mejorar la previsión para mejorar la capacidad de una organización para predecir eventos futuros en procesos comerciales clave, como ventas y servicios. Para lograr este objetivo, se utilizó investigación relacionada con la tecnología de minería de datos, que analiza la información mediante un enfoque predictivo. El desarrollo de la investigación incluye el desarrollo de diversos modelos utilizando métodos como regresión, redes neuronales y árboles de decisión basados en la base de datos histórica de organizaciones automovilísticas, preselección de datos utilizando métodos como matrices de correlación y PCA (Análisis de Componentes Principales). Finalmente, luego de evaluar los resultados obtenidos al comparar los modelos propuestos, encontramos que mejores resultados para el pronóstico de ventas se obtuvieron con el modelo de red neuronal implementado usando PCA, mientras que el modelo de servicio de mantenimiento principal se implementó con el modelo de bosque aleatorio, por lo que en la comparación Los métodos de árbol de decisión fueron bosques aleatorios y se utilizan algoritmos de árbol de clasificación. Para mejorar los resultados del modelo, se realizó un ajuste de hiperparámetros en el primer modelo para lograr una mejor precisión del modelo (95,31%). Por otro lado, en el segundo modelo se cortó el árbol, lo que mejoró la precisión (95,14%). Sin embargo, en comparación con los dos últimos modelos, el bosque aleatorio con ajuste de hiperparámetros funcionó mejor. Por tanto, se puede concluir que este es el mejor método.

Según Ccopa y Chávez (2015) en su trabajo de investigación titulado “Modelo Predictivo Basado en Minería de Datos Para la Mejora en la Toma de decisiones del departamento de Cirugía del Hospital Regional Manuel Núñez Butrón” tuvo por objetivo general desarrollar un Modelo Predictivo basado en

Minería de Datos para la mejora en la toma de decisiones del Departamento Cirugía del Hospital Regional Manuel Núñez Butrón. El problema que se descubrió fue la falta de información necesaria para lograr una adecuada toma de decisiones más allá de los procesos de obtención de esa información. Este estudio toma la cirugía como ejemplo. Además de la imposibilidad de obtener predicciones de las tarifas hospitalarias, también existen problemas importantes con los registros de datos escritos a mano, lo que dificulta la identificación de los datos. El criterio de éxito determinado en este estudio es el control de los indicadores hospitalarios, que es la base para una adecuada toma de decisiones, para lo cual se aplicó la técnica de regresión lineal de exploración de datos- regresión del modelo de pronóstico. El método DM recopila datos utilizando el software y los sistemas de información Weka. Finalmente, se realizaron pruebas de hipótesis y se demostró empíricamente que "el desarrollo de modelos predictivos basados en minería de datos puede mejorar la toma de decisiones, lo cual es aceptable ya que aumentó un 91,97%". Las ecuaciones se calcularon a partir de datos de un hospital y se utilizaron técnicas de regresión lineal. La precisión de las sugerencias Se determinó que el modelo es efectivo con una precisión del 83,9% de los datos del modelo.

Según Saldaña (2015) en su trabajo de investigación titulado "Modelo Predictivo de minería de datos de apoyo a la gestión hospitalaria sobre la morbilidad de pacientes Hospitalizados" La minería de datos en el sector salud permite identificar los perfiles de salud en los pacientes, ayuda a comprender el patrón de comportamiento, a través del historial de información almacenada que hace posible su gestión cotidiana, siendo así esta información diversa y compleja. Este trabajo de investigación propone el uso de un marco estándar para actividades de minería de datos para construir un modelo de pronóstico para respaldar la morbilidad de los pacientes hospitalizados basado en un algoritmo de análisis de series de tiempo del modelo ARIMA (promedio móvil

integrado autorregresivo). Por Box y Jenkis (Box G.E.P. & Jenkins, 1973) Contiene información histórica sobre los pacientes del Hospital LaGuardia Victor Ramos durante los últimos 7 años. Este estudio utilizó la metodología CRISP-DM (Proceso estándar de descubrimiento de datos entre industrias), que incluye comprensión empresarial, comprensión de datos, preparación, modelado, evaluación e implementación de datos. Por lo tanto, se llevaron a cabo la recuperación de datos, la conversión de datos, la carga de datos, la limpieza de datos, el diseño de la ejecución de datos "HEALTHMINING", la selección de variables como datos de entrada y la construcción de modelos para luego construir un modelo de pronóstico que nos permita comprender la morbilidad entre los pacientes ingresados en el hospital GAL durante los próximos tres años. Por lo que con base en estos resultados, asumiendo 2 indicadores que cumplan con las reglas de inferencia establecidas, se formuló la hipótesis: "La creación de un modelo de investigación de datos predictivos para el GAL Hospital Huaraz apoyará el manejo de la morbilidad hospitalaria en este hospital".

### **2.1.3. A nivel Local**

En este caso en la búsqueda del repositorio local no se encontró temas similares al trabajo de investigación.

## **2.2. Bases teóricas – científicas**

### **2.2.1. Minería de datos**

Las definiciones de minería de datos son:

Gabriel Amorocho define que: "Es el proceso de recolectar y analizar grandes cantidades de datos (típicamente históricos) desde diferentes perspectivas para identificar patrones, correlaciones o tendencias ocultas entre muchas variables para posteriormente inferir las razones subyacentes para estos comportamientos y proyectarlos al futuro" (Amorocho, 2015)

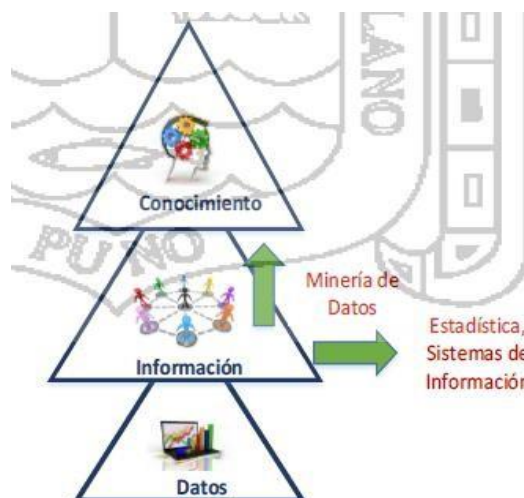
Según Palomino Mariana Garrido es: “El uso de la tecnología para descubrir información útil, aplicable e importante en grandes volúmenes de datos”. Esta definición se aplica a varios contextos empresariales y se puede reformular como “el conjunto de métodos, que junto con un profundo conocimiento del negocio, están orientados a identificar, en grandes volúmenes de datos, relaciones y tendencias ocultas hasta el momento” (Garrido, 2012).

Carlos Márquez Vera menciona: “La minería de datos está orientada a trabajar con cantidades muy grandes de datos (millones y billones de datos). En cambio, la estadística no suele funcionar tan bien en bases de datos de tan gran tamaño y alta dimensionalidad” (Márquez, 2015)

Luis Carlos Molina Felix dice: “Es un conjunto de áreas (Estadística, inteligencia artificial, reconocimiento de patrones, computación gráfica y bases de datos) que tienen como propósito la identificación de conocimiento obtenido a partir de las bases de datos que aporten un sesgo o apoyo hacia la toma de decisión” (Molina, 2014).

De las definiciones anteriores concluimos que la minería de datos es el procesamiento de grandes cantidades de datos generando como resultado el conocimiento, ver Figura 1”.

**Figura 1.**  
*Definición de Minería de Datos.*



## 2.2.2. Modelos de Minería de Datos

*Figura 2.*  
*Fases de la Minería de datos*



## 2.2.3. Modelo Predictivo

Oldemar Rodríguez Rojas define modelo predictivo como: “entrenar modelos que se supervisa se dirige a algo para detectar algo o predecir.

El modelo predictivo es también conocido como Aprendizaje Supervisado, Método Predictivo, Estadistical learning, Analytics” (Rodríguez Rojas, Metodos Predictivos en Minería de Datos - PERSON, 2015).

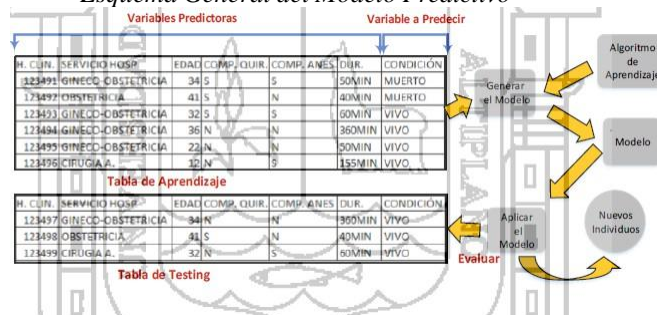
Sandra Mendoza y Fernando Henríquez resaltan que: “el modelo predictivo se emplea para estimar valores futuros de variables de interés. El proceso se basa en la información histórica de los datos mediante las cuales se predice un comportamiento de los datos, ya sea mediante clasificaciones, categorizaciones o regresiones. Los modelos predictivos siguen un aprendizaje supervisado, que consiste en aprender mediante el control de un supervisor o maestro que determina la respuesta que se desea generar del sistema. El atributo a predecir se conoce como variable dependiente u objetivo, mientras que los atributos utilizados para realizar la predicción se llaman variables independientes o de exploración” (Mendoza Castillo, Sandra Jaqueline; Zavaleta Henríquez, Fernando, 2013).



José Mondragón resalta que los modelos predictivos: “son parte fundamental de la minería de datos y por ello siempre es necesario revisar su desempeño y poder de predicción para ser mejorados o inclusive cambiados en sus parámetros ya que los patrones de comportamiento pueden variar según la época o momento en el que fueron desarrollados” (Mondragón, 2012).

Álvaro Jiménez y Hugo Álvarez destacan que: “en predicción, el objetivo es desarrollar un modelo que pueda inferir una variable a partir de alguna combinación de otras variables incluidas en los datos. La predicción requiere etiquetas para la variable de salida para un conjunto de datos limitado, donde una etiqueta suponga una información fiable sobre el valor de la variable de salida en casos específicos. De todas maneras, en algunos casos es importante considerar el grado en el que estas etiquetas puedan ser aproximadas o inciertas. En un segundo tipo de uso, los métodos de predicción son utilizados para predecir cuál será el valor de salida en contextos donde no es deseable obtener una etiqueta para una construcción (por ejemplo, en ocasiones en las que no haya datos etiquetados)”. (Jiménez Galindo, Álvaro; Álvarez García, Hugo, 2010). En la Figura 3 se muestra el esquema general de modelo predictivo.

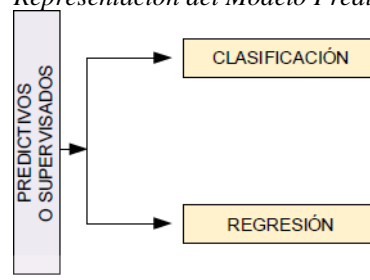
**Figura 3.**  
*Esquema General del Modelo Predictivo*



#### 2.2.4. Métodos del Modelo Predictivo

Según Gloria Hernández Gómez: “Los modelos de predicción se pueden dividir en dos métodos: clasificación y regresión”. En la figura 4 se muestra una representación del modelo de predicción.

**Figura 4.**  
*Representación del Modelo Predictivo.*



##### - **Clasificación:**

Oldemar Rodríguez Rojas dice que: “el método de clasificación se representa de la siguiente manera que: dada una colección de registros o conjunto de entrenamiento cada registro contiene un conjunto de variables que son los atributos denominadas  $x$  también llamadas variables predictoras, con una variable adicional que es la clase denominada  $y$  o variable a predecir” (Rodríguez Rojas, 2015).

Oldemar Rodríguez Rojas también destaca que: “el objetivo de la Clasificación que es encontrar un modelo (una función o algoritmo) para predecir la clase a la que pertenecía cada registro, esta asignación de una clase se debe hacer con la mayor precisión posible. Un conjunto de prueba (tabla de testing) se utiliza para determinar la precisión del modelo. Por lo general, el conjunto de datos dado se divide en dos conjuntos al azar: el de entrenamiento y el de prueba” (Rodríguez Rojas, 2015)

Oldemar Rodríguez define clasificación: “en una forma matemática dada una base de datos  $D=\{t_1,t_2,t_3,\dots,t_n\}$  de tuplas o registros (individuos) y un conjunto de clases  $C=\{C_1,C_2,C,\dots,C_m\}$ , el problema de la clasificación es encontrar una función  $f:D\rightarrow C$  tal que cada  $t_i$  es asignada una clase  $C_j$ .”

La función  $f:D\rightarrow C$  podría ser una Red Neuronal, una Regresión Lineal, un árbol de Decisión y otros” (Rodríguez Rojas, 2015).

#### - **Regresión**

Jay Devore rescata que: La palabra "regresión" fue utilizada por primera vez como término estadístico en 1877 por Sir Francis Galton, quien realizó un estudio que demostró que la altura de los niños nacidos de padres altos tendía a retroceder o volver a la altura promedio de la población. Designó la palabra regresión como el nombre del proceso general de predecir una variable (la estatura de los niños) a partir de otra (la estatura del padre o de la madre). Más tarde, los estadísticos acuñaron el término regresión múltiple para describir el proceso mediante el cual se utilizan varias variables para predecir otra” (Devore, 2008)

Héctor Rodríguez, destaca que: “La regresión ayuda a revelar la dependencia de los valores de los atributos de otros atributos en la misma entidad u objeto. La regresión en la clasificación de datos es similar a los árboles de decisión, pero predice atributos continuos en lugar de atributos discretos”.

En muchos estudios de minería de datos estadísticos destinados a la toma de decisiones profesionales o personales, uno de los principales objetivos es establecer relaciones que permitan predecir una o más variables en función de otras variables.

#### **2.2.5. Técnicas del Modelo Predictivo.**

Wilson Andrés Castillo Rojas Resalta que: “las siguientes Técnicas son utilizados en el Modelo Predictivo” (Castillo Rojas, 2013):

- Redes Neuronales.
- Árboles de Decisión (ID3,C4.5.C5.0).
- Árboles de Decisión (CART).

- Otros Árboles de Decisión.
- Regresión Lineal y Logarítmica.
- Regresión Logística.
- Naive Bayes.
- Vecinos Más Próximos.
- Algoritmos Genéticos y Evolutivos.
- Máquinas de Vectores de Soporte.
- CN2 Reglas.
- Análisis Discriminante Multivariante

### 2.3. Definición de términos básicos

**Árboles de Decisión:** Es una técnica útil para problemas en los que se presentan decisiones secuenciales. Aunque esta técnica es de mayor utilidad para situaciones en que el riesgo está presente también es empleada en condiciones de certeza (Liz, 2012).

**Casos de Uso:** Es una descripción de la secuencia de interacciones que se producen entre un actor y el sistema, cuando el actor usa el sistema para llevar a cabo una tarea específica. (Ferré Grau, Xavier; Sánchez Segura, María Isabel, 2011).

**Dato:** Es la unidad mínima de la información. Es un elemento aislado, recabado para un cierto fin, pero que no ha pasado por un proceso que lo interrelacione con otros de manera funcional para el fin previsto. Los datos recogen un conjunto de hechos (Molina López, 2006).

**Diagramas de Secuencia:** Resaltan la ordenación temporal de los mensajes que se intercambian (Booch, Grady; Rumbaugh, jim; Jacobson, Ivar, 2013).

**Información:** Es un conjunto de mecanismos que permiten a un individuo estructurarlos de una manera determinada. Los patrones, asociaciones o relaciones entre los datos proporcionan información (Calderón Méndez, 2006).

**Método:** Es una serie de pasos sucesivos, conducen a una meta. El objetivo del profesionalista es llegar a tomar las decisiones y una teoría que permita generalizar y resolver de la misma forma problemas semejantes en el futuro (Eduardo, 2015)

**Predicción:** Es un pronóstico razonable y verificable, sobre un hecho o acontecimiento nuevo o desconocido. Se caracteriza por anticipar lo que va a ocurrir; la inferencia, en cambio, explica o interpreta lo que ya ha ocurrido (Jankovic, 2011).

**Técnicas:** Son los procedimientos e instrumentos que se utilizan para acceder al conocimiento (Morone, 2013).

**UML:** Lenguaje Unificado de Modelado proporciona un vocabulario y unas reglas para permitir una comunicación. En este caso, este lenguaje se centra en la representación gráfica de un sistema. (Hernández, 2011).

## **2.4. Formulación de Hipótesis**

### **2.4.1. Hipótesis general**

El modelo predictivo basado en Minería de datos Mejora en la Toma de Decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023.

### **2.4.2. Hipótesis Específicas**

- a) Aplicando el modelo predictivo mejora la toma de decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023.
- b) Evaluando el modelo predictivo mejora la toma de decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023.

## **2.5. Identificación de Variables**

### **2.5.1. Variables independientes**

Modelo Predictivo.

## 2.5.2. Variables dependientes

Toma de decisiones.

## 2.6. Definición Operacional de variables e indicadores

*Tabla 1. Definición Operacional de Variables*

<b>Variable</b>	<b>Definición conceptual</b>	<b>Dimensiones</b>
INDEPENDIENTE Modelo Predictivo	También conocidos como modelos de predicción, son un conjunto de herramientas y técnicas estadísticas que sirven para pronosticar y predecir el comportamiento ante un evento. Y tienen el objetivo de predecir y pronosticar resultados probables a futuro	- Desarrollo de modelo predictivo. - Evaluación
DEPENDIENTE Toma de decisiones	La toma de decisiones es el proceso mediante el cual se realiza una elección entre alternativas o formas de resolver diferentes situaciones de la vida. Ciertos factores pueden influir en las decisiones que tomamos diariamente, a través de algunas técnicas y actividades prácticas	-Proceso

**Fuente:** Elaboración propia

## **CAPITULO III**

### **METODOLOGÍA Y TECNICAS DE INVESTIGACIÓN**

#### **3.1. Tipo de Investigación**

El tipo de la investigación del presente estudio es aplicada. Según Ortega (2017), es el estudio científico que busca resolver un problema o planteamiento específico, que se caracteriza por buscar la aplicación o utilización de los conocimientos que se adquieren.

#### **3.2. Nivel de investigación**

El presente trabajo de investigación utilizó el nivel cuantitativo, porque se basa en el estudio y análisis de la realidad a través de diferentes procedimientos basados en la medición.

#### **3.3. Métodos de investigación**

La investigación que realizaré es la inducción analítica, la cual se encarga de encontrar las causas de los hechos estableciendo relaciones causales. En este sentido, la investigación explicativa puede ocuparse tanto de la identificación de causas como de la identificación de causas (investigación

post facto), como de los efectos (investigación experimental), mediante la prueba de hipótesis. Sus resultados y conclusiones constituyen el nivel más profundo de conocimientos” según (Hernández Sampieri et al., 2018)

### **3.4. Diseño de investigación**

La investigación de acuerdo a las características del problema, objetivo y la hipótesis es de tipo experimental, de diseño pre experimental, con pre-test y post-test, los diseños pre experimentales no presentan grupo control.

El grupo experimental estuvo conformado por los datos de los pacientes registrados en el periodo 2022-2023, en el Hospital Daniel Alcides Carrión.

La representación gráfica es la siguiente:

*G1: 01 X 02*

Dónde:

*G1: Grupo experimental.*

*X: Tratamiento con el modelo de minería de datos.*

*01: Test antes del experimento.*

*02: Test después del experimento.*

Este diseño con grupo experimental permitió la comparación de resultados pre test y post test, con un alto grado de probabilidad, ante la situación con modelo de minería de datos y sin modelo (variable independiente), ha sido un factor determinante para la toma de decisiones en la predicción de datos de la Medicina rehabilitación del Hospital Daniel Alcides Carrión (Variable dependiente).

### **3.5. Población y muestra**

#### **3.5.1. Población**

Según Hernández Sampieri (2018) Este es un grupo de personas que forman parte del departamento y nos apoyan en la obtención de pruebas y resultados.



Todos los datos de pacientes atendidos 2022 y 2023 del Hospital Daniel Alcides Carrión Pasco.

### **3.5.2. Muestra**

Para Arias, Villasis y Miranda (2016) comento que “la muestra es una parte de la población que es característico de un todo y se opta por adquirir información acerca de las variables”.

Se toma como muestra 137 pacientes en medicina rehabilitación.

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

Se utilizaron fichas de observación y observación directa.

### **3.7. Selección, validación y confiabilidad de los instrumentos de investigación**

#### **Fichas de observación:**

Se obtuvo de la base de datos en Microsoft Excel los registros están realizados los años 2022 y 2023 se encuentra en el anexo.

### **3.8. Técnicas de procesamiento y análisis de datos.**

Se utilizó la técnica de regresión lineal.

### **3.9. Tratamiento Estadístico**

Se aplicaron las fases de la metodología CRISP–DM:

- Comprensión del negocio.
- Comprensión de data.
- Preparación de data.
- Modelado.
- Evaluación.

### **3.10. Orientación ética filosófica y epistémica.**

En este estudio se consideran los siguientes factores:

Mantendremos el formulario de registro confidencial y garantizaremos el anonimato, por lo que sus datos personales serán omitidos.

## **CAPITULO IV**

### **RESULTADOS Y DISCUSIÓN**

#### **4.1. Descripción del trabajo de campo**

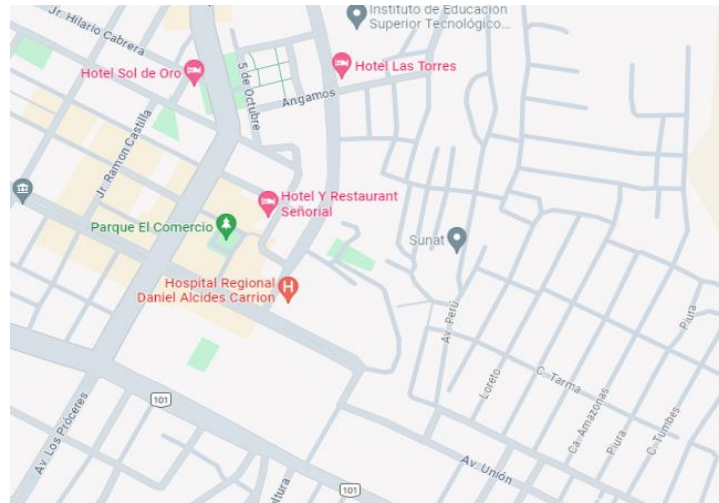
##### **Misión del Hospital Daniel Alcides Carrión**

La Misión del Hospital es que somos un Hospital de referencia Regional perteneciente al ministerio de Salud que brinda servicios de Salud Integral y Especializada, comprometidos en el bienestar de la población pasqueña.

##### **Visión del Hospital Daniel Alcides Carrión**

Ser un Hospital líder en la Región, en la prestación de servicios de salud integral especializada, que considera el recurso humano lo más valioso de su organización y brindar a la población el acceso a los servicios de salud con calidad.

**Figura 5.**  
**Ubicación.**



## 4.2. Presentación, análisis e interpretación de resultados

### 4.2.1. Comprensión de Negocio

Según el organigrama del consultorio de Medicina de rehabilitación, la máxima autoridad es el director del hospital, seguido por el técnico administrativos y entre otras divisiones según corresponda, los objetivos del departamento son brindar rehabilitación a lo siguiente:

**Tabla 2.**

*Diagnóstico para rehabilitación*

Distensión Muscular
Espolón Calcáneo
Gota no especificada
Hemiplejia Espástica
Lumbago con ciática
Luxación de la articulación del hombro
Trastornos de los meniscos
Parálisis cerebral
Parálisis de Bell
Pie plano

Procedimiento terapéutico
Retardo de desarrollo
Tendinitis
Terapia de aprendizaje
Trastorno de disco cervical

#### 4.2.2. Comprensión de data

##### Recolección de data y descripción

La primera etapa de metodología CRISP-DM describe la alineación de los objetivos del proyecto con la investigación, por ende, el levantamiento de la información fue recolectada mediante la base de datos del Hospital Daniel Alcides Carrión que comprende del periodo 2022 al 2023, dicha información se alinea con los objetivos del proyecto y la información adquirida servirá para analizarla y poder interpretar el modelo que se desea diseñar.

La información recolectada comprende los siguientes indicadores:

*Tabla 3.*

*Diccionario de datos*

Variable	Descripción
IdCita	Código de número de cita
Año	En qué año fue atendido
Mes	En qué mes fue atendido
UPS	Código de consultorio de rehabilitación
IdPaciente	Identificación del paciente
Sexo	Sexo de paciente
IdPersonal	Identificación del personal
Personal	Nombre del personal que fue atendido
Financiador	Modalidad de atención
IdTurno	Turno de atención
DiagnosticoProcedimiento	Diagnóstico de enfermedad

#### 4.2.3. Preparación de data

Para preparar la data se lleva una descripción estadística de todas las etiquetas para determinar la media y desviación estándar, también se muestran

los gráficos de la data para verificar la igualdad de partición en la clase a predecir, un factor clave en esta fase es el sobre muestreo en la clase minoritaria y se identifica en la observación de la data.

**Figura 6.**  
*Datos de pacientes en el Consultorio de Rehabilitación*

	A	B	C	D	E	F	G	H	I	J	K
	IdCita	Año	Mes	UPS	IdPaciente	Sexo	IdPersonal	Personal	Financiad	IdTurno	DIAGNOSTICO PROCEDIMIENTO
1	794307903	2022	12	MEDICINA REHABILITACION	32471967979	M	26947676979	REQUIZ REYE S.I.S	M	M	EVALUACION DE TERAPIA OCUPACIONAL
2	794308097	2022	12	MEDICINA REHABILITACION	45662811979	F	26947676979	REQUIZ REYE USUARIO	M	M	TECNICAS DE INTEGRACION SENSORIAL PARA MEJORAR EL PRO
3	794308348	2022	12	MEDICINA REHABILITACION	43534406979	M	26947676979	REQUIZ REYE S.I.S	M	M	ACTIVIDADES TERAPEUTICAS CONTACTO DIRECTO (UNO A UNO
4	794308543	2022	12	MEDICINA REHABILITACION	12326852979	F	26947676979	REQUIZ REYE S.I.S	M	M	EVALUACION DE TERAPIA OCUPACIONAL
5	794308781	2022	12	MEDICINA REHABILITACION	12292765979	M	26947676979	REQUIZ REYE S.I.S	M	M	EVALUACION DE TERAPIA OCUPACIONAL
6	794315172	2022	12	MEDICINA REHABILITACION	35665230979	M	26947676979	REQUIZ REYE S.I.S	T	M	EVALUACION DE TERAPIA OCUPACIONAL
7	794315520	2022	12	MEDICINA REHABILITACION	32430962979	M	26947676979	REQUIZ REYE USUARIO	T	M	TECNICAS DE INTEGRACION SENSORIAL PARA MEJORAR EL PRO
8	794316011	2022	12	MEDICINA REHABILITACION	41168420979	M	26947676979	REQUIZ REYE USUARIO	T	M	TECNICAS DE INTEGRACION SENSORIAL PARA MEJORAR EL PRO
9	794643463	2022	12	MEDICINA REHABILITACION	14662144979	M	20308398979	MEZA ALIAGI S.I.S	M	M	DISARTRIA Y ANARTRIA
10	794643463	2022	12	MEDICINA REHABILITACION	14662144979	M	20308398979	MEZA ALIAGI S.I.S	M	M	TERAPIA DE LENGUAJE/SESION
11	794643463	2022	12	MEDICINA REHABILITACION	14662144979	M	20308398979	MEZA ALIAGI S.I.S	M	M	TERAPIA DEL COMPLEJO OROFACIAL/SESION
12	794643463	2022	12	MEDICINA REHABILITACION	14662144979	M	20308398979	MEZA ALIAGI S.I.S	M	M	TECNICAS DE TERAPIA MANUAL
13	794644041	2022	12	MEDICINA REHABILITACION	43534406979	M	20308398979	MEZA ALIAGI S.I.S	M	M	TRASTORNO DEL LENGUAJE EXPRESIVO
14	794644041	2022	12	MEDICINA REHABILITACION	43534406979	M	20308398979	MEZA ALIAGI S.I.S	M	M	TERAPIA DE LENGUAJE/SESION
15	794644041	2022	12	MEDICINA REHABILITACION	43534406979	M	20308398979	MEZA ALIAGI S.I.S	M	M	TERAPIA DEL COMPLEJO OROFACIAL/SESION
16	794644041	2022	12	MEDICINA REHABILITACION	43534406979	M	20308398979	MEZA ALIAGI S.I.S	M	M	TECNICAS DE TERAPIA MANUAL
17	794644041	2022	12	MEDICINA REHABILITACION	43534406979	M	20308398979	MEZA ALIAGI S.I.S	M	M	DESARROLLO DE HABILIDADES COGNITIVAS PARA MEJORAR LA
18	794644711	2022	12	MEDICINA REHABILITACION	10371083979	M	20308398979	MEZA ALIAGI S.I.S	M	M	TRASTORNO DEL LENGUAJE EXPRESIVO
19	794644711	2022	12	MEDICINA REHABILITACION	10371083979	M	20308398979	MEZA ALIAGI S.I.S	M	M	TERAPIA DE LENGUAJE/SESION
20	794644711	2022	12	MEDICINA REHABILITACION	10371083979	M	20308398979	MEZA ALIAGI S.I.S	M	M	TECNICAS DE TERAPIA MANUAL
21	794644711	2022	12	MEDICINA REHABILITACION	10371083979	M	20308398979	MEZA ALIAGI S.I.S	M	M	DESARROLLO DE HABILIDADES COGNITIVAS PARA MEJORAR LA
22	794644711	2022	12	MEDICINA REHABILITACION	10371083979	M	20308398979	MEZA ALIAGI S.I.S	M	M	TERAPIA DE APRENDIZAJE

Parte de la preparación es ver la distribución de la etiqueta predictiva, una forma de realizar es graficar si la clase está balanceada debido que un desbalanceo no permite realizar una predicción deseada.

Preparación de datos al Software SPSS v26.

**Figura 7.**  
*Datos de pacientes en el Consultorio de Rehabilitación en SPSS*

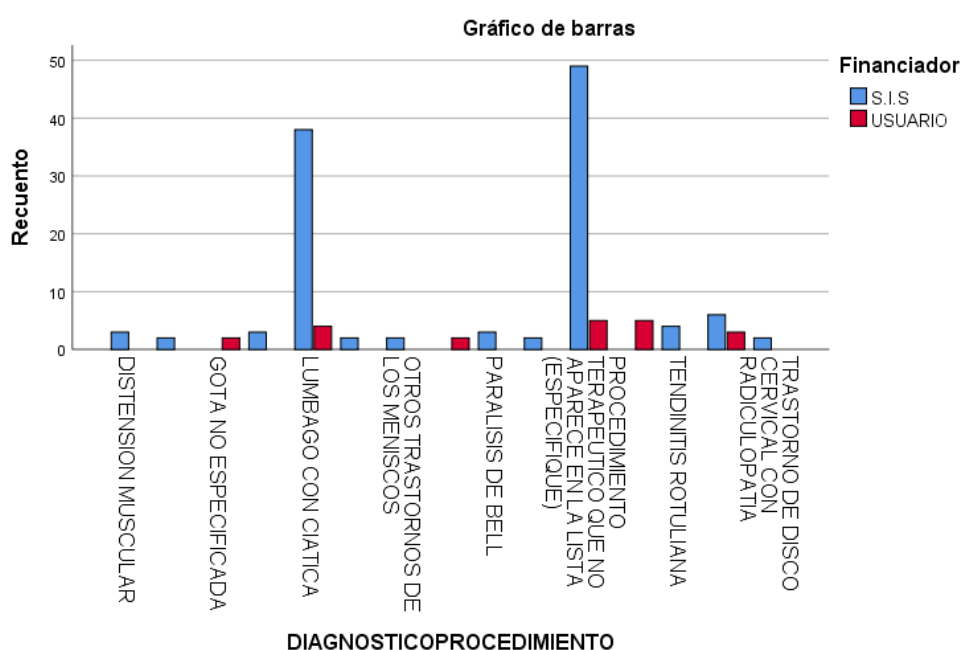
	IdCita	Año	Mes	UPS	IdPacient	Sexo	IdPersona	Personal	Financiad	IdTurno	
1	794307903	2022	12	MEDICINA REHABILITACION	32471967979	M	26947676979	REQUIZ REYES YAMELIT LIDIA	S I S	M	EVALUACION DE TERAPIA C
2	794308097	2022	12	MEDICINA REHABILITACION	45662811979	F	26947676979	REQUIZ REYES YAMELIT LIDIA	USUARIO	M	TECNICAS DE INTEGRACION
3	794308348	2022	12	MEDICINA REHABILITACION	43534406979	M	26947676979	REQUIZ REYES YAMELIT LIDIA	S I S	M	ACTIVIDADES TERAPEUTICA
4	794308543	2022	12	MEDICINA REHABILITACION	12326852979	F	26947676979	REQUIZ REYES YAMELIT LIDIA	S I S	M	EVALUACION DE TERAPIA C
5	794308781	2022	12	MEDICINA REHABILITACION	12292765979	M	26947676979	REQUIZ REYES YAMELIT LIDIA	S I S	M	EVALUACION DE TERAPIA C
6	794315172	2022	12	MEDICINA REHABILITACION	35665230979	M	26947676979	REQUIZ REYES YAMELIT LIDIA	S I S	T	EVALUACION DE TERAPIA C
7	794315520	2022	12	MEDICINA REHABILITACION	32430962979	M	26947676979	REQUIZ REYES YAMELIT LIDIA	USUARIO	T	TECNICAS DE INTEGRACION
8	794316011	2022	12	MEDICINA REHABILITACION	41168420979	M	26947676979	REQUIZ REYES YAMELIT LIDIA	USUARIO	T	TECNICAS DE INTEGRACION
9	794643463	2022	12	MEDICINA REHABILITACION	14662144979	M	20308398979	MEZA ALIAGA CINTHIA VANESSA	S I S	M	DISARTRIA Y ANARTRIA
10	794643463	2022	12	MEDICINA REHABILITACION	14662144979	M	20308398979	MEZA ALIAGA CINTHIA VANESSA	S I S	M	TERAPIA DE LENGUAJE/SEE
11	794643463	2022	12	MEDICINA REHABILITACION	14662144979	M	20308398979	MEZA ALIAGA CINTHIA VANESSA	S I S	M	TERAPIA DEL COMPLEJO OF
12	794643463	2022	12	MEDICINA REHABILITACION	14662144979	M	20308398979	MEZA ALIAGA CINTHIA VANESSA	S I S	M	TECNICAS DE TERAPIA MAN
13	794644041	2022	12	MEDICINA REHABILITACION	43534406979	M	20308398979	MEZA ALIAGA CINTHIA VANESSA	S I S	M	TRASTORNO DEL LENGUAJE
14	794644041	2022	12	MEDICINA REHABILITACION	43534406979	M	20308398979	MEZA ALIAGA CINTHIA VANESSA	S I S	M	TERAPIA DE LENGUAJE/SEE
15	794644041	2022	12	MEDICINA REHABILITACION	43534406979	M	20308398979	MEZA ALIAGA CINTHIA VANESSA	S I S	M	TECNICAS DE TERAPIA MAN
16	794644041	2022	12	MEDICINA REHABILITACION	43534406979	M	20308398979	MEZA ALIAGA CINTHIA VANESSA	S I S	M	TECNICAS DE TERAPIA MAN
17	794644041	2022	12	MEDICINA REHABILITACION	43534406979	M	20308398979	MEZA ALIAGA CINTHIA VANESSA	S I S	M	DESARROLLO DE HABILIDA
18	794644711	2022	12	MEDICINA REHABILITACION	10371083979	M	20308398979	MEZA ALIAGA CINTHIA VANESSA	S I S	M	TRASTORNO DEL LENGUAJE
19	794644711	2022	12	MEDICINA REHABILITACION	10371083979	M	20308398979	MEZA ALIAGA CINTHIA VANESSA	S I S	M	TERAPIA DE LENGUAJE/SEE
20	794644711	2022	12	MEDICINA REHABILITACION	10371083979	M	20308398979	MEZA ALIAGA CINTHIA VANESSA	S I S	M	TECNICAS DE TERAPIA MAN
21	794644711	2022	12	MEDICINA REHABILITACION	10371083979	M	20308398979	MEZA ALIAGA CINTHIA VANESSA	S I S	M	DESARROLLO DE HABILIDA

**Figura 8.**  
Análisis descriptivo de pacientes en el Consultorio de Rehabilitación en SPSS.

**Tabla cruzada DIAGNOSTICOPROCEDIMIENTO\*Financiador**

Recuento		Financiador		Total
		S.I.S	USUARIO	
DIAGNOSTICOPROCEDIMIENTO	DISTENSION MUSCULAR	3	0	3
	ESPOLON CALCANEEO	2	0	2
	GOTA NO ESPECIFICADA	0	2	2
	HEMIPLEJIA ESPASTICA	3	0	3
	LUMBAGO CON CIATICA	38	4	42
	LUXACION DE LA ARTICULACION DEL HOMBRO	2	0	2
	OTROS TRASTORNOS DE LOS MENISCOS	2	0	2
	PARALISIS CEREBRAL ESPASTICA	0	2	2
	PARALISIS DE BELL	3	0	3
	PIE PLANO	2	0	2
	PROCEDIMIENTO TERAPEUTICO QUE NO APARECE EN LA LISTA (ESPECIFIQUE)	49	5	54
	RETARDO DEL DESARROLLO	0	5	5
	TENDINITIS ROTULIANA	4	0	4
	TERAPIA DE APRENDIZAJE	6	3	9
	TRASTORNO DE DISCO CERVICAL CON RADICULOPATIA	2	0	2
<b>Total</b>		<b>116</b>	<b>21</b>	<b>137</b>

**Figura 9.**  
Gráfico de barras Análisis descriptivo de pacientes en el Consultorio de Rehabilitación en SPSS.



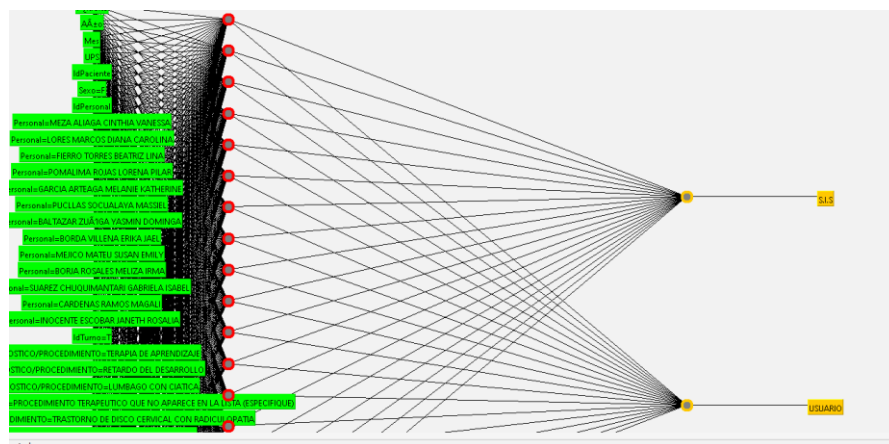
**Interpretación:** El 42% de los pacientes en consultorio de Medicina de Rehabilitación son diagnosticados con procedimiento terapéutico que no aparece en la lista quiere decir que son luxaciones en el cuerpo.

#### 4.2.4. Modelado.

Construcción del Modelo Perceptrón Multicapa, se grafica con 3 capas.

- La primera capa corresponde a los valores de entradas que son las 12 variables analizadas y procesadas en las fases de la metodología CRISP DM.
- La segunda capa corresponde a la capa oculta, que es parte del proceso en las redes neuronales para realizar una retropropagación y el modelo sea óptimo, esta capa consta de 2 neuronas.
- Por último, la capa deseada a predecir es de 1 neurona cuyo valor es 0,1.

*Figura 10.  
Red MLP.*



La imagen describe, un modelo secuencial de un clasificador por capa, la función Dense describe las variables de 1 capa, en la primera capa describe el valor 'units' como las neuronas de la primera capa, la capa de entrada consta de 11 etiquetas, la inicialización de pesos mediante la derivada 'kernel' es de rango de -0.5 a +0.5 debido que presenta mayor valor de error cuadrático medio y es óptimo para realizar una convergencia.

#### 4.2.5. Evaluación.

##### Matriz de clasificación de MLP

Una forma de validar modelos de precisión es por medio de la matriz de clasificación y éstas se determinan por las siguientes fórmulas:

La precisión.

$$Precisión_i = \frac{N_{ii}}{\sum_{k=1}^n N_{ki}}$$

Recall.

$$Recall_i = \frac{N_{ii}}{\sum_{k=1}^n N_{ik}}$$

F1 Score.

$$F - Score_i = \frac{2 \times Precision_i * Recall_i}{Precision_i + Recall_i}$$

El modelo obtuvo el siguiente reporte de métricas:

**Figura 11.**  
*Reporte de Métricas.*

```
Time taken to build model: 0.92 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      137          100   %
Incorrectly Classified Instances    0             0   %
Kappa statistic                    1
Mean absolute error                 0.005
Root mean squared error             0.0121
Relative absolute error             1.0009 %
Root relative squared error         2.4329 %
Total Number of Instances          137

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
1.000  0.000  1.000  1.000  1.000  1.000  1.000  1.000  Procedimiento (CPT)
1.000  0.000  1.000  1.000  1.000  1.000  1.000  1.000  Diagnostico (CIEX)
Weighted Avg.  1.000  0.000  1.000  1.000  1.000  1.000  1.000  1.000

=== Confusion Matrix ===

 a  b  <-- classified as
63  0  | a = Procedimiento (CPT)
 0 74  | b = Diagnostico (CIEX)
```



## Construcción del Modelo Naive Bayes

Para hacer la predicción se necesita calcular la probabilidad que hay en una instancia de datos que pertenezca cada variable o clase. El proceso sigue los siguientes 4 pasos: Calcular la función de la densidad de probabilidad gaussiana, las probabilidades de las variables, predicción y precisión de la estimación:

```
Time taken to build model: 1.05 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      115          83.9416 %
Incorrectly Classified Instances    22          16.0584 %
Kappa statistic                    0.3563
Mean absolute error                 0.1902
Root mean squared error            0.3627
Relative absolute error             72.1956 %
Root relative squared error        100.6287 %
Total Number of Instances          137

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
                0.914   0.571   0.898     0.914   0.906     0.357   0.676   0.896   S.I.S
                0.429   0.086   0.474     0.429   0.450     0.357   0.676   0.446   USUARIO
Weighted Avg.   0.839   0.497   0.833     0.839   0.836     0.357   0.676   0.827

=== Confusion Matrix ===

  a  b  <-- classified as
106 10 |  a = S.I.S
 12  9 |  b = USUARIO
```

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

Donde:

$P(A/B)$ : la probabilidad de ocurrencia del evento A, dado el evento B, ya ha ocurrido.

$P(A)$  - Probabilidad de ocurrencia del evento A.

$P(B)$  - Probabilidad de ocurrencia del evento B.

$P(B/A)$  - Probabilidad de ocurrencia del evento B, dado que el evento A ya ha ocurrido.

**Figura 12.**  
**Reporte de Naive Bayes**

```

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      132          96.3504 %
Incorrectly Classified Instances     5           3.6496 %
Kappa statistic                     0.927
Mean absolute error                  0.0648
Root mean squared error              0.1642
Relative absolute error              13.0442 %
Root relative squared error          32.9255 %
Total Number of Instances           137

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
1.000  0.068  0.926  1.000  0.962  0.929  0.999  0.999  Procedimiento (CPT)
0.932  0.000  1.000  0.932  0.965  0.929  0.999  0.999  Diagnostico (CIEIX)
Weighted Avg.  0.964  0.031  0.966  0.964  0.964  0.929  0.999  0.999

=== Confusion Matrix ===

 a  b  <-- classified as
63  0 | a = Procedimiento (CPT)
 5 69 | b = Diagnostico (CIEIX)

Time taken to build model: 0.01 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      112          81.7518 %
Incorrectly Classified Instances     25          18.2482 %
Kappa statistic                     0.3359
Mean absolute error                  0.2149
Root mean squared error              0.3859
Relative absolute error              81.5614 %
Root relative squared error          107.0611 %
Total Number of Instances           137

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
0.879  0.524  0.903  0.879  0.891  0.337  0.642  0.892  S.I.S
0.476  0.121  0.417  0.476  0.444  0.337  0.642  0.376  USUARIO
Weighted Avg.  0.818  0.462  0.828  0.818  0.822  0.337  0.642  0.813

=== Confusion Matrix ===

 a  b  <-- classified as
102 14 | a = S.I.S
 11 10 | b = USUARIO

```

La precisión de predicción de Naive Bayes alcanzó el 86%, siendo la predicción de la presente investigación, de una total de 137 pacientes.

### Construcción del Modelo Árbol de decisión

Extracción de las variables más redundantes aplicando “Feature Importance”, que es el grado de importancia de las variables para construir un modeloprobabilístico, el cual es una clasificación artificial donde el número de variables (n\_features) son la cantidad de los valores de entradas al árbol y el

gráfico arroja como resultado la importancia de las características informativas de (n\_informative).

**Figura 13.**  
*Reporte de Árbol de decisión*

```

Time taken to build model: 0.01 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      137          100  %
Incorrectly Classified Instances    0             0  %
Kappa statistic                     1
Mean absolute error                 0
Root mean squared error             0
Relative absolute error             0  %
Root relative squared error         0  %
Total Number of Instances          137

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
                1.000    0.000    1.000     1.000    1.000     1.000    1.000    1.000    Procedimiento (CPT)
                1.000    0.000    1.000     1.000    1.000     1.000    1.000    1.000    Diagnostico (CIEX)
Weighted Avg.   1.000    0.000    1.000     1.000    1.000     1.000    1.000    1.000

=== Confusion Matrix ===

 a  b  <-- classified as
63  0  | a = Procedimiento (CPT)
 0 74 | b = Diagnostico (CIEX)

```

A continuación, se puede apreciar la importancia de cada una de las variables de la data de experimentación en la presente investigación, también se puede expresar como variables más influyentes según el cálculo con la técnica de árbol de decisión J48, que a la vez alcanzó un 100% de precisión, estos resultados son representación de influencia de manera general, sin especificar las clases que podría tener un problema.

**Figura 14.**  
Árbol de decisión



### Evaluación de precisión de las técnicas

En el presente proyecto de minería de datos se experimentó tres principales algoritmos, Redes neuronales perceptrón multicapa, Naive Bayes y

árbol de decisión, entre los cuales la técnica árbol de decisión alcanzó a la mejor precisión.

**Tabla 4.**

*Resultado de las técnicas de estudio*

Ítem	Técnica	Precisión
1	Perceptrón Multicapa	83%
2	Naive Bayes	86%
3	Árbol de decisión	100%

### 4.3. Prueba de Hipótesis

#### 4.3.1. Hipótesis General

Para probar la hipótesis, se utilizó una comparación de los datos reales y los datos devueltos por el modelo y se promedió de acuerdo con la siguiente fórmula.

##### A. Predicción

$$Precision = \frac{N^{\circ} \text{ de predicciones correctas}}{N^{\circ} \text{ total de predicciones}} \times 100$$

**Tabla 5.**  
*Tabla de predicción*

<b>Indicadores de diagnóstico</b>	<b>Precisión Con Modelo</b>	<b>Precisión Sin Modelo</b>	<b>Incorrecto</b>	<b>% de Precisión</b>
Distensión Muscular	137	137	0	100
Espolón Calcáneo	137	120	16	94.11
Gota no especificada	137	67	117	85
Hemiplejia Espástica	137	137	0	100
Lumbago con ciática	137	137	0	100
Luxación de la articulación del hombro	137	137	0	100
Trastornos de los meniscos	137	100	37	27
Parálisis cerebral	137	137	0	100
Parálisis de Bell	137	120	17	12
Pie plano	137	125	12	8
Procedimiento terapéutico	137	134	3	2.18
Retardo de desarrollo	137	136	1	0.72
<b>Total</b>	<b>1644</b>	<b>1487</b>	<b>203</b>	<b>84%</b>

## **B. Datos históricos**

$$DH = \frac{\Sigma (\text{indicadores con modelo} - \Sigma \text{indicadores sin modelo})}{\Sigma \text{indicadores sin modelo}} \times 100$$

<b>Indicadores</b>	<b>% Datos Históricos Sin Modelo</b>	<b>% Datos Históricos Con Modelo</b>
Distensión Muscular	SI	SI
Espolón Calcáneo	SI	SI
Gota no especificada	SI	SI
Hemiplejia Espástica	NO	SI
Lumbago con ciática	SI	SI
Luxación de la articulación del hombro	SI	SI
Trastornos de los meniscos	NO	SI
Parálisis cerebral	NO	SI
Parálisis de Bell	NO	SI
Pie plano	NO	SI
Procedimiento terapéutico	NO	SI
Retardo de desarrollo	NO	SI
<b>TOTAL</b>	<b>5</b>	<b>12</b>

### **C. Mejora en la toma de decisiones**

$$MTD = \frac{Precisión + LH}{2} \times 100$$

$$MTD = \frac{84+100}{2} \times 100 = 92\%$$

Entonces, los modelos de minería de datos son buenos porque mejoran la toma de decisiones en un 92% y el grado está en el rango del 76-100%.

#### **4.4. Discusión de resultados**

Mediante el uso de minería de datos, la Oficina de Medicina de Rehabilitación del Hospital Daniel Alcides Carrión Pasco ha desarrollado un modelo predictivo que ha resultado en una mejora en la toma de decisiones en un 92%.

Las redes neuronales de perceptrón multicapa, Naive Bayes y Decision Tree fueron los tres algoritmos principales probados en este proyecto de minería de datos, entre los cuales la técnica árbol de decisión alcanzó a la mejor precisión: En el análisis de Perceptrón Multicapa con un 83%, Naive Bayes 86% y Árbol de decisión 100%.

Según (Ccopa Mamani & Chavez Viza, 2015) en los resultado obtenidos de su trabajo “Se desarrolló el Modelo Predictivo Basado en Minería de Datos para la Mejora en la Toma de Decisiones del Departamento de Cirugía del Hospital Regional Manuel Nuñez Butrón, gracias al promedio del Coeficiente de Correlación de los indicadores hospitalarios de 81.34% que se obtuvo la evaluación del modelo es correcta mejorando así en la toma de decisiones en un porcentaje de 91.97%”.



## CONCLUSIONES

El modelo de minería de datos mejora en la toma de decisiones en un porcentaje del 92%, con un grado excelente ya que está en el rango de 76-100%.

Se aplicó la metodología de CRISP-DM para:

- Utilizar tres algoritmos de predicción, a saber, perceptrón multicapa, árbol de decisión y Naive Bayes, para clasificar a los pacientes de la Clínica de Medicina de Rehabilitación del Hospital Daniel Alcides Carrión Pasco
- En la Oficina de Medicina de Rehabilitación, se utilizó la técnica del árbol de decisión J48 para clasificar a los pacientes según sus variables, siendo el diagnóstico, la financiación y la precisión del 100% los factores más importantes
- La validación cruzada del modelo arrojó el siguiente informe de métricas: Perceptrón multicapa (83%), Árbol de decisión (100%) y Naive Bayes (86%), siendo este último el más preciso.

## **RECOMENDACIONES**

- Incorporar el modelo de minería de datos predictivo en futuras iniciativas de gestión hospitalaria, considerando la creación de software para la recopilación de información y la mejora de la capacidad de toma de decisiones a través de informes automatizados.
- Al utilizar técnicas de extracción de datos, se sugiere mejorar el proceso de toma de decisiones de los centros de atención en otros centros de rehabilitación y comparar sus resultados con los de diferentes instituciones.
- Utilizar esta investigación como referencia para futuros estudios sobre este tema.

## REFERENCIAS BIBLIOGRÁFICAS

- Amoroch, G. (2015). *Minería de datos: que es y 5 consejos para aprovecharla en las grandes empresas*.
- Becerra, J., & Enrique, V. (2021). *Data Mining para modelo predictivo de ventas y servicios de mantenimiento en un concesionario Automotriz Ligero*. Universidad de Lima.
- Calderón Méndez, N. (2006). *Minería de datos una Herramienta para la Toma de Decisiones*.
- Castillo Rojas, W. A. (2013). *Inteligencia de Negocios. Minería de Datos III. Chile*.
- Ccopa Mamani, M., & Chavez Viza, S. V. (2015). Modelo Predictivo Basado en Minería de Datos Para la Mejora en la Toma de decisiones del departamento de Cirugía del Hospital Regional Manuel Núñez Butrón [UNIVERSIDAD NACIONAL DEL ALTIPLANO]. In *Universidad Nacional del Altiplano*.  
<http://repositorio.unap.edu.pe/handle/UNAP/1911>
- Devore, J. (2008). *Probabilidad y Estadística para Ingenierías y Ciencias*. México: Cengage Learning Editores.
- Eduardo, A. (2015). *Definición General del Metodo*.
- García, F. (2013). *Aplicación de técnicas de Minería de Datos a datos obtenidos por el Centro Andaluz de Medio Ambiente (CEAMA)*. UNIVERSIDAD GRANADA.
- Garrido, P. M. (2012). *Búsqueda de patrones en Bases de Datos y su aplicación en las PYMES*.
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2018). *Metodología de la Investigación* (S. A. D. C. . McGRAW-HILL / INTERAMERICANA EDITORES (ed.); Sexta).

Ministerio de Salud. (2013). Resolución Ministerial N° 827-2013-MINSA. NTS N° 105-MINSA/DGSP-V.01 “Norma técnica de salud para la atención integral de Salud Materna.” In *Norma Técnica-Minsa* (p. 61).

Saldaña, E. (2015). *Modelo Predictivo de minería de datos de apoyo a la gestión hospitalaria sobre la morbilidad de pacientes Hospitalizados*. UNIVERSIDAD PRIVADA ANTENOR ORREGO.

Uvidia Fassler, M. I., Cisneros Barahona, A. S., Méndez Naranjo, P. M., & Villa Yáñez, H. M. (2019). Minería de datos para la toma de decisiones en la unidad de nivelación y admisión universitaria ecuatoriana [UNIVERSIDAD ECUATORIANA]. In *Cumbres* (Vol. 4, Issue 2). <https://doi.org/10.48190/cumbres.v4n2a5>

**ANEXOS**

# Instrumento de Recolección de datos

## Fichas de Registro en Base de datos

DATA DICIEMBRE2022 - Excel (Error de activación de productos)

Inicio    Insertar    Disposición de página    Fórmulas    Datos    Revisar    Vista    Programador    Ayuda    Acrobat    ¿Qué desea hacer?

Calibri 11 A A<sup>+</sup> A<sup>-</sup>    General    Formato condicional    Dar formato como tabla    Estilos de celda    Insertar Eliminar Formato    Ordenar y filtrar    Buscar y seleccionar    Edición

Portapapeles    Fuente    Alineación    Número    Celdas

AE1    MesActualPaciente

	A	B	C	D	E	F	G	H	I	J	K	L	M
	IdCita	Año	Mes	Dia	FechaAtencion	Lote	NumPag	NumReg	IdUps	UPS	IdEstablecimiento	IdPaciente	Paciente
1	794307903	2022	12	1	44896 NF1		20	5	302201	MEDICINA REHABILITACION	979	32471967979	DE LA TORRE SOTO ANDER
2	794308097	2022	12	1	44896 NF1		20	6	302201	MEDICINA REHABILITACION	979	45662811979	CONDOR CASTRO RN MUJE
3	794308348	2022	12	1	44896 NF1		20	7	302201	MEDICINA REHABILITACION	979	43534406979	ALBINO CONDOR JORDAN
4	794308543	2022	12	1	44896 NF1		20	8	302201	MEDICINA REHABILITACION	979	12326852979	MEZA VEGA SONIA
5	794308781	2022	12	1	44896 NF1		20	9	302201	MEDICINA REHABILITACION	979	12292765979	LEDESMA CASIMIRO ATILIC
6	794315172	2022	12	1	44896 NF2		20	1	302201	MEDICINA REHABILITACION	979	35665230979	RAMIREZ FERNANDEZ SELI
7	794315250	2022	12	1	44896 NF2		20	2	302201	MEDICINA REHABILITACION	979	32430962979	CASTA?EDA TAIZE YADIEL H
8	794316011	2022	12	1	44896 NF2		20	3	302201	MEDICINA REHABILITACION	979	41168420979	VIVAS CONDEZO CHRISTOF
9	794643463	2022	12	1	44896 CM1		1	1	302201	MEDICINA REHABILITACION	979	14662144979	RECUAY CASTRO MARCO A
10	794643463	2022	12	1	44896 CM1		1	1	302201	MEDICINA REHABILITACION	979	14662144979	RECUAY CASTRO MARCO A
11	794643463	2022	12	1	44896 CM1		1	1	302201	MEDICINA REHABILITACION	979	14662144979	RECUAY CASTRO MARCO A
12	794643463	2022	12	1	44896 CM1		1	1	302201	MEDICINA REHABILITACION	979	14662144979	RECUAY CASTRO MARCO A
13	794643463	2022	12	1	44896 CM1		1	1	302201	MEDICINA REHABILITACION	979	14662144979	RECUAY CASTRO MARCO A
14	794644041	2022	12	1	44896 CM1		1	2	302201	MEDICINA REHABILITACION	979	43534406979	ALBINO CONDOR JORDAN
15	794644041	2022	12	1	44896 CM1		1	2	302201	MEDICINA REHABILITACION	979	43534406979	ALBINO CONDOR JORDAN
16	794644041	2022	12	1	44896 CM1		1	2	302201	MEDICINA REHABILITACION	979	43534406979	ALBINO CONDOR JORDAN
17	794644041	2022	12	1	44896 CM1		1	2	302201	MEDICINA REHABILITACION	979	43534406979	ALBINO CONDOR JORDAN
18	794644041	2022	12	1	44896 CM1		1	2	302201	MEDICINA REHABILITACION	979	43534406979	ALBINO CONDOR JORDAN
19	794644711	2022	12	1	44896 CM1		1	3	302201	MEDICINA REHABILITACION	979	10371083979	USCUCHAGUA ROJAS ADRI
20	794644711	2022	12	1	44896 CM1		1	3	302201	MEDICINA REHABILITACION	979	10371083979	USCUCHAGUA ROJAS ADRI
21	794644711	2022	12	1	44896 CM1		1	3	302201	MEDICINA REHABILITACION	979	10371083979	USCUCHAGUA ROJAS ADRI
22	794644711	2022	12	1	44896 CM1		1	3	302201	MEDICINA REHABILITACION	979	10371083979	USCUCHAGUA ROJAS ADRI
23	794644711	2022	12	1	44896 CM1		1	3	302201	MEDICINA REHABILITACION	979	10371083979	USCUCHAGUA ROJAS ADRI

DATA DICIEMBRE2022 - Excel (Error de activación de productos)

Inicio    Insertar    Disposición de página    Fórmulas    Datos    Revisar    Vista    Programador    Ayuda    Acrobat    ¿Qué desea hacer?

Calibri 11 A A<sup>+</sup> A<sup>-</sup>    General    Formato condicional    Dar formato como tabla    Estilos de celda    Insertar Eliminar Formato    Ordenar y filtrar    Buscar y seleccionar    Edición

Portapapeles    Fuente    Alineación    Número    Celdas

AE1    MesActualPaciente

	A	B	C	D	E	F	G	H	I	J	K	L	M
	IdCita	Año	Mes	Dia	FechaAtencion	Lote	NumPag	NumReg	IdUps	UPS	IdEstablecimiento	IdPaciente	Paciente
76	795537137	2022	12	1	44896 D05		1	5	302201	MEDICINA REHABILITACION	979	12309482979	ROMERO BEZADA JUAN MA
77	795537137	2022	12	1	44896 D05		1	5	302201	MEDICINA REHABILITACION	979	12309482979	ROMERO BEZADA JUAN MA
78	795537137	2022	12	1	44896 D05		1	5	302201	MEDICINA REHABILITACION	979	12309482979	ROMERO BEZADA JUAN MA
79	795537137	2022	12	1	44896 D05		1	5	302201	MEDICINA REHABILITACION	979	12309482979	ROMERO BEZADA JUAN MA
80	795537227	2022	12	1	44896 D05		1	6	302201	MEDICINA REHABILITACION	979	12269319979	CAPCHA DAVIRAN VILMA L
81	795537227	2022	12	1	44896 D05		1	6	302201	MEDICINA REHABILITACION	979	12269319979	CAPCHA DAVIRAN VILMA L
82	795537227	2022	12	1	44896 D05		1	6	302201	MEDICINA REHABILITACION	979	12269319979	CAPCHA DAVIRAN VILMA L
83	795537227	2022	12	1	44896 D05		1	6	302201	MEDICINA REHABILITACION	979	12269319979	CAPCHA DAVIRAN VILMA L
84	795537227	2022	12	1	44896 D05		1	6	302201	MEDICINA REHABILITACION	979	12269319979	CAPCHA DAVIRAN VILMA L
85	795537227	2022	12	1	44896 D05		1	6	302201	MEDICINA REHABILITACION	979	12269319979	CAPCHA DAVIRAN VILMA L
86	795537334	2022	12	1	44896 D05		1	7	302201	MEDICINA REHABILITACION	979	22038805979	HUAYLLA AYALA RUTH ESTI
87	795537334	2022	12	1	44896 D05		1	7	302201	MEDICINA REHABILITACION	979	22038805979	HUAYLLA AYALA RUTH ESTI
88	795537334	2022	12	1	44896 D05		1	7	302201	MEDICINA REHABILITACION	979	22038805979	HUAYLLA AYALA RUTH ESTI
89	795537334	2022	12	1	44896 D05		1	7	302201	MEDICINA REHABILITACION	979	22038805979	HUAYLLA AYALA RUTH ESTI
90	795537334	2022	12	1	44896 D05		1	7	302201	MEDICINA REHABILITACION	979	22038805979	HUAYLLA AYALA RUTH ESTI
91	795537334	2022	12	1	44896 D05		1	7	302201	MEDICINA REHABILITACION	979	22038805979	HUAYLLA AYALA RUTH ESTI
92	795537398	2022	12	1	44896 D05		1	8	302201	MEDICINA REHABILITACION	979	4094039979	PALACIN SINCHE JAIR JHAS
93	795537398	2022	12	1	44896 D05		1	8	302201	MEDICINA REHABILITACION	979	4094039979	PALACIN SINCHE JAIR JHAS
94	795537398	2022	12	1	44896 D05		1	8	302201	MEDICINA REHABILITACION	979	4094039979	PALACIN SINCHE JAIR JHAS
95	795537398	2022	12	1	44896 D05		1	8	302201	MEDICINA REHABILITACION	979	4094039979	PALACIN SINCHE JAIR JHAS
96	795537398	2022	12	1	44896 D05		1	8	302201	MEDICINA REHABILITACION	979	4094039979	PALACIN SINCHE JAIR JHAS
97	795537398	2022	12	1	44896 D05		1	8	302201	MEDICINA REHABILITACION	979	4094039979	PALACIN SINCHE JAIR JHAS
98	795537476	2022	12	1	44896 D05		1	9	302201	MEDICINA REHABILITACION	979	12170173979	HEREDIA ARAGON DE ATE

Excel window: DATA DICIEMBRE2022 - Excel (Error de activación de productos)

File menu: Archivo, Inicio, Insertar, Disposición de página, Fórmulas, Datos, Revisar, Vista, Programador, Ayuda, Acrobat, ¿Qué desea hacer?, Compartir

Font settings: Calibri, 11, Fuente, Alineación, Combinar y centrar

Number settings: General, Número, Formato condicional, Estilos de celdas

Layout settings: Insertar, Eliminar, Formato, Celdas

Ordering and search: Ordenar y filtrar, Buscar y seleccionar, Edición

	A	B	C	D	E	F	G	H	I	J	K	L	M
3	800633253	2022	12	1	44896 GAB		1	11	302201	MEDICINA REHABILITACION		979	19264979979 CAÑA CAJAHUAMAN WILL
4	800633253	2022	12	1	44896 GAB		1	11	302201	MEDICINA REHABILITACION		979	19264979979 CAÑA CAJAHUAMAN WILL
5	800633366	2022	12	1	44896 DI2		1	3	302201	MEDICINA REHABILITACION		979	10923426979 LAUREANO TACURI RONAL
6	800633366	2022	12	1	44896 DI2		1	3	302201	MEDICINA REHABILITACION		979	10923426979 LAUREANO TACURI RONAL
7	800633366	2022	12	1	44896 DI2		1	3	302201	MEDICINA REHABILITACION		979	10923426979 LAUREANO TACURI RONAL
8	800633366	2022	12	1	44896 DI2		1	3	302201	MEDICINA REHABILITACION		979	10923426979 LAUREANO TACURI RONAL
9	800633445	2022	12	1	44896 MB		1	5	302201	MEDICINA REHABILITACION		979	12119469979 JESUS SANTOS LUZMILA
10	800633445	2022	12	1	44896 MB		1	5	302201	MEDICINA REHABILITACION		979	12119469979 JESUS SANTOS LUZMILA
11	800633445	2022	12	1	44896 MB		1	5	302201	MEDICINA REHABILITACION		979	12119469979 JESUS SANTOS LUZMILA
12	800633445	2022	12	1	44896 MB		1	5	302201	MEDICINA REHABILITACION		979	12119469979 JESUS SANTOS LUZMILA
13	800633445	2022	12	1	44896 MB		1	5	302201	MEDICINA REHABILITACION		979	12119469979 JESUS SANTOS LUZMILA
14	800633445	2022	12	1	44896 MB		1	5	302201	MEDICINA REHABILITACION		979	12119469979 JESUS SANTOS LUZMILA
15	800633517	2022	12	1	44896 GAB		1	12	302201	MEDICINA REHABILITACION		979	12178589979 CAMONES BUSTILLOS JAVI
16	800633517	2022	12	1	44896 GAB		1	12	302201	MEDICINA REHABILITACION		979	12178589979 CAMONES BUSTILLOS JAVI
17	800633517	2022	12	1	44896 GAB		1	12	302201	MEDICINA REHABILITACION		979	12178589979 CAMONES BUSTILLOS JAVI
18	800633517	2022	12	1	44896 GAB		1	12	302201	MEDICINA REHABILITACION		979	12178589979 CAMONES BUSTILLOS JAVI
19	800633517	2022	12	1	44896 GAB		1	12	302201	MEDICINA REHABILITACION		979	12178589979 CAMONES BUSTILLOS JAVI
20	800633773	2022	12	1	44896 MB		1	6	302201	MEDICINA REHABILITACION		979	12173904979 RIVERA ESPINOZA NILDA M
21	800633773	2022	12	1	44896 MB		1	6	302201	MEDICINA REHABILITACION		979	12173904979 RIVERA ESPINOZA NILDA M
22	800633773	2022	12	1	44896 MB		1	6	302201	MEDICINA REHABILITACION		979	12173904979 RIVERA ESPINOZA NILDA M
23	800633773	2022	12	1	44896 MB		1	6	302201	MEDICINA REHABILITACION		979	12173904979 RIVERA ESPINOZA NILDA M
24	800633773	2022	12	1	44896 MB		1	6	302201	MEDICINA REHABILITACION		979	12173904979 RIVERA ESPINOZA NILDA M
25	800633773	2022	12	1	44896 MB		1	6	302201	MEDICINA REHABILITACION		979	12173904979 RIVERA ESPINOZA NILDA M

## Matriz de Consistencia

**Tema:** “Modelo Predictivo basado en minería de datos para la mejora en toma de decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023”

PROBLEMA GENERAL	OBJETIVO GENERAL	HIPÓTESIS GENERAL	VARIABLE INDEPENDIENTE	DIMENSIÓN	DISEÑO	POBLACIÓN Y MUESTRA
¿En qué medida el Modelo Predictivo basado en Minería de Datos mejora la toma de decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023?	Desarrollar un modelo predictivo basado en minería de datos para la mejora de toma de decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023.	El modelo predictivo basado en Minería de datos Mejora en la Toma de Decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023	Modelo Predictivo.	- Desarrollo de modelo predictivo. - Evaluación	<b>Diseño:</b> Pre - Experimental  <b>Tipo de Investigación</b> Aplicada	<b>POBLACIÓN</b> Todos los datos de pacientes atendidos 2022 del Hospital Daniel Alcides Carrión Pasco.  <b>MUESTRA</b> Se toma como muestra 137 pacientes en medicina rehabilitación.
PROBLEMA ESPECÍFICO	OBJETIVO ESPECÍFICO	HIPÓTESIS ESPECÍFICA	VARIABLE DEPENDIENTE	DIMENSIÓN	MÉTODO DE INVESTIGACIÓN	TÉCNICAS - INSTRUMENTOS
¿Se podrá desarrollar el modelo predictivo para la mejora en toma de decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023?	Desarrollar el modelo predictivo para la mejora en toma de decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023.	Aplicando el modelo predictivo mejora la toma de decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023.	Toma de decisiones.	- Proceso	<b>Método</b> Analítica, inductiva <b>Enfoque</b> Cuantitativo	<b>Técnicas:</b>  - Fichas.
¿Se podrá evaluar el modelo predictivo para la mejora en toma de decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023?	Evaluar el modelo predictivo para la mejora en toma de decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023.	Evaluando el modelo predictivo mejora la toma de decisiones en el Consultorio de Medicina rehabilitación en el Hospital Daniel Alcides Carrión Pasco, 2023.				