



UNIVERSIDAD CÉSAR VALLEJO

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

Sistema de análisis predictivo para la gestión de ventas
de la empresa Tauro A&F SAC

**TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE:
Ingeniero de Sistemas**

AUTOR(ES):

Aragon Farfan, Christian Jesus Junior (orcid.org/0000-0001-9410-5564)

Sanchez Altamirano, Cleiver Dermalí (orcid.org/0000-0003-2349-8916)

ASESOR:

Dr. Mendoza Apaza, Fernando (orcid.org/0000-0001-7981-8291)

LÍNEA DE INVESTIGACIÓN:

Sistemas de Información y Comunicaciones

LÍNEA DE RESPONSABILIDAD SOCIAL UNIVERSITARIA:

Desarrollo económico, empleo y emprendimiento

LIMA - PERÚ
2022

Dedicatoria

A la familia que me motivó para seguir enfocándonos en el término de la carrera y seguir creciendo profesionalmente.

Aragón Farfan Christian Jesús Junior

Agradecimiento

En especial a la familia para poder seguir en el término carrera profesional y en el asesor por el apoyo brindado en el desarrollo del presente proyecto de investigación.

Índice de Contenidos

Carátula	i
Dedicatoria.....	ii
Agradecimiento.....	iii
Índice de contenidos.	iv
Índice de tablas	v
Índice de gráficos y figuras.....	vi
Resumen.....	vii
Abstract.....	viii
I. INTRODUCCIÓN	1
II. MARCO TEÓRICO.....	4
III. METODOLOGÍA.....	10
3.1 Tipo y diseño de investigación.....	10
3.2. Variables y operacionalización	10
3.3. Población, muestra y muestreo	11
3.4. Técnicas e instrumentos de recolección de datos	12
3.5. Procedimientos.....	12
3.6. Método de análisis de datos	12
3.7. Aspectos éticos	13
IV. RESULTADOS	14
V. DISCUSIÓN.....	23
VI. CONCLUSIONES	25
VII. RECOMENDACIONES.....	26
REFERENCIAS.....	27

Índice de tablas

Tabla 1. Medidas descriptivas de los indicadores	14
Tabla 2. Prueba de normalidad del índice de rotación de productos	16
Tabla 3. Prueba Wilcoxon del indicador índice de rotación de productos	17
Tabla 4. Prueba Z del indicador índice de rotación de productos	17
Tabla 5. Prueba de normalidad del índice de ganancia de productos	18
Tabla 6. Prueba Wilcoxon del índice de ganancia de productos	18
Tabla 7. Prueba Z del índice de ganancia de productos	19
Tabla 8. Prueba de normalidad del índice de precisión de productos	20
Tabla 9. Prueba Wilcoxon del índice de precisión de productos	20
Tabla 10. Prueba Z del índice de precisión de productos	21
Tabla 11. Recursos tecnológicos para el desarrollo del modelo	55
Tabla 12. Campos del reporte de ventas	56
Tabla 13. Descripción de los datos	56
Tabla 14. Selección de datos	57

Índice de gráficos y figuras

Figura 1. Ingresos de la organización	2
Figura 2. Cantidad de Productos Vendidos	2
Figura 3. Fórmulas para el cálculo de la muestra	10
Figura 4. Pretest y postest del índice de rotación de productos	15
Figura 5. Pretest y postest del índice de ganancia de productos	15
Figura 6. Pretest y postest del índice de precisión de venta	16
Figura 7. Lenguaje de programación para el desarrollo del modelo	53
Figura 8. Visual Studio Code, editor de código	53
Figura 9. Google Colaboratory	54
Figura 10. Verificación de la calidad de los datos	57
Figura 11. Limpieza de los datos	58
Figura 12. Estructuración de los datos	58
Figura 13. Formateo de los datos	58
Figura 14. Datos de entrada	59
Figura 15. Técnicas de modelado	60
Figura 16. Librerías del modelo	60
Figura 17. Lectura de los datos	60
Figura 18. Lectura de las columnas	61
Figura 19. Renombramiento de las columnas	61
Figura 20. Identificación del índice de las columnas	61
Figura 21. Muestra de datos del dataframe	61
Figura 22. Rango de fechas del dataset	62
Figura 23. Validación del producto a predecir	62
Figura 24. Comportamiento del producto a predecir	63
Figura 25. Columnas como entrada al modelo	64
Figura 26. Entrenamiento del modelo	64
Figura 27. Función para usar el modelo	64
Figura 28. Pruebas del modelo	65
Figura 29. Exportación del modelo	66
Figura 30. Pantalla de inicio del aplicativo	67
Figura 31. Filtros de la predicción	67
Figura 32. Botón para realizar la predicción	68
Figura 33. Botón para descargar la predicción	68
Figura 34. Archivo con la predicción	68

Resumen

El objetivo de la investigación fue determinar la influencia del sistema de análisis predictivo para la gestión de ventas de la empresa TAURO A&F SAC.

La investigación desarrollada fue de tipo aplicada, de enfoque cuantitativo y de diseño pre-experimental. El instrumento empleado fue la ficha de registro y se empleó la metodología CRISP-DM para el desarrollo del proyecto. Los resultados luego de la implementación nos permitieron concluir que el sistema de análisis predictivo mejoró la gestión de ventas en la empresa TAURO A&F SAC; mejorando el índice de rotación de productos en 2%, el índice de ganancia de productos en 15% y el índice de precisión de productos en 9% de la gestión de ventas.

Palabras clave: Machine Learning, proceso de ventas y productos

Abstract

The objective of the investigation was to determine the influence of the predictive analysis system for the sales management of the company TAURO A&F SAC.

The research carried out was of an applied type, with a quantitative approach and a pre-experimental design. The instrument used was the registration form and the CRISP-DM methodology was used for the development of the project. The results after the implementation allowed us to conclude that the predictive analysis system improved sales management in the company TAURO A&F SAC; improving product turnover rate by 2%, product win rate by 15% and product accuracy rate by 9% of sales management.

Keywords: Machine Learning, sales process and products

I. INTRODUCCIÓN

En esta era donde la tecnología enfocada en datos está generando gran relevancia en las organizaciones a nivel estratégico, operacional, social, laboral, entre otros.

Saber cómo aprovechar las distintas tecnologías para apoyar las decisiones y obtener una ventaja frente a los competidores, que nos permita prever situaciones y poder decidir lo más conveniente a la organización.

En el ámbito internacional, en un artículo publicado en The New York Times, Lohr (2018, p.1), menciona que la inteligencia artificial se encarga de identificar patrones y realizar predicciones en el área en la cual se implemente tomando como ejemplo Amazon en la que mantiene constante aprendizaje sobre el comportamiento de compra de sus consumidores estando a la vanguardia en sus necesidades.

En el ámbito nacional, en un artículo publicado en Ernst & Young Global Ltd, Escudero (2021, p.40), menciona que la tendencia en el uso de la automatización robótica e inteligencia artificial tiene un 65% de consideración en aplicarla y, el área de operaciones y cadena de suministro tiene un 27% donde se considera que se pueden aplicar.

El siguiente informe se desarrolla en la empresa Tauro A&F SAC la cual se enfoca en el rubro de venta, mantenimiento y reparación de motocicletas. El área de ventas está encargada de atender a los clientes mayoristas y minoristas por medio de un sistema de ventas transaccional, el área de almacén es el que se encuentra saturado de productos ya que lo que se hace es comprar apenas las empresas presenten promociones para posteriormente venderlos en un tiempo indeterminado teniendo una baja rotación de productos y porque las compras se hacen pensando que se ahorran costos al comprar cuando se ofertan promociones además del dinero que se encuentra invertido pero que no genera retorno a corto plazo pero sin tener en consideración si la cantidad de productos que se compra es la adecuada.

Actualmente la organización no tiene una herramienta que los pueda ayudar a ser más eficiente en la futura demanda de productos. Lo que se tiene planeado para resolver esta problemática es aplicar machine learning y generar un modelo que pueda hacer más eficiente y productiva el área de ventas.

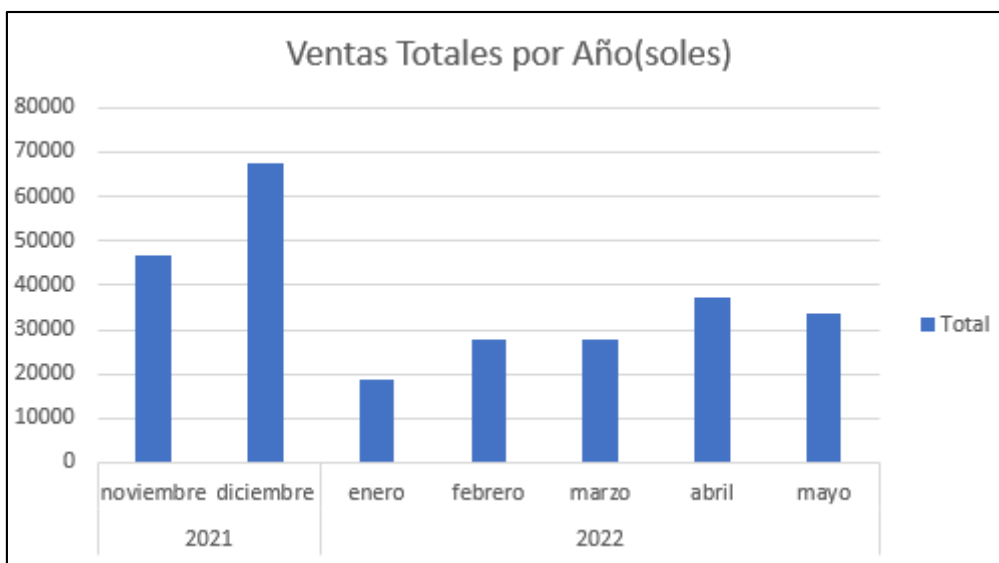


Figura 1: Ingresos de la organización (Elaboración propia)

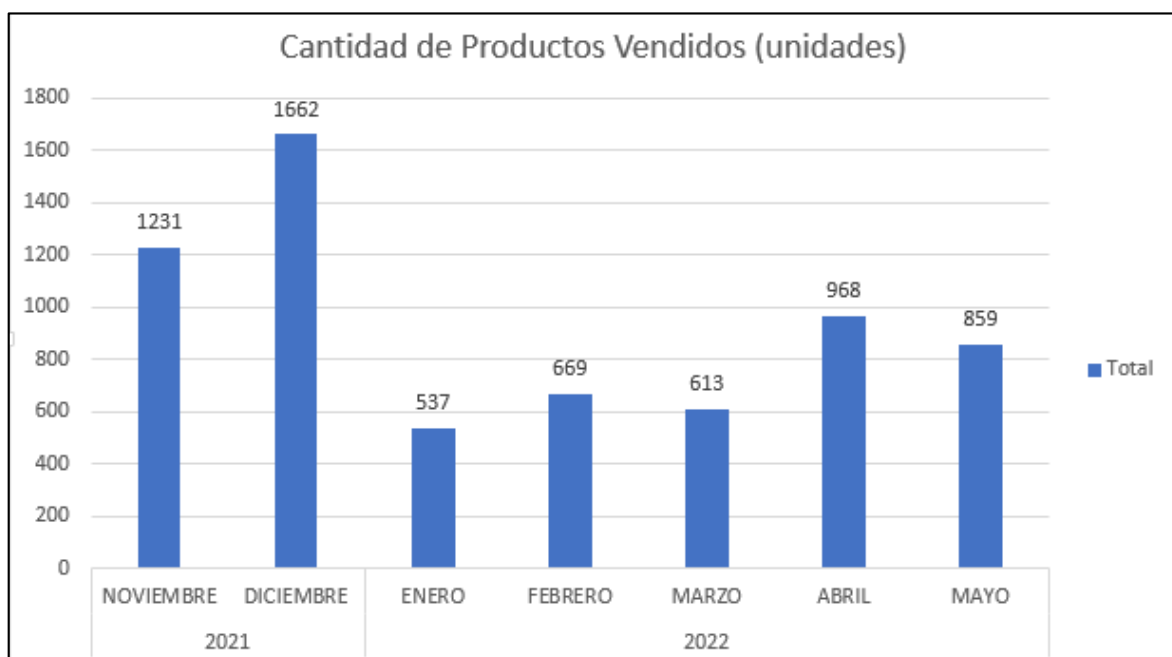


Figura 2: Cantidad de Productos Vendidos (Elaboración propia)

Teniendo en consideración la realidad del problema, se propone como problema general ¿Cómo influye el sistema de análisis predictivo para la gestión de ventas de la empresa TAURO A&F SAC?, y los siguientes problemas específicos:

PE1: ¿Cómo influye el sistema de análisis predictivo en incrementar el índice de rotación de productos para la gestión de ventas de la empresa TAURO A&F SAC?

PE2: ¿Cómo influye el sistema de análisis predictivo en el índice de ganancia de los productos para la gestión de ventas de la empresa TAURO A&F SAC?

PE3: ¿Cómo influye el sistema de análisis predictivo en incrementar el índice de precisión de productos para la gestión de ventas de la empresa TAURO A&F SAC?

Además se planteó el objetivo general que fue determinar la influencia del sistema de análisis predictivo para la gestión de ventas de la empresa TAURO A&F SAC, y los siguientes objetivos específicos:

OE1: Determinar la influencia del sistema de análisis predictivo para incrementar el índice de rotación de productos en la gestión de ventas de la empresa TAURO A&F SAC.

OE2: Determinar la influencia del sistema de análisis predictivo para incrementar el índice de ganancia de los productos en la gestión de ventas de la empresa TAURO A&F SAC

OE3: Determinar la influencia del sistema de análisis predictivo para incrementar el índice de precisión de productos en la gestión de ventas de la empresa TAURO A&F SAC.

De acuerdo a los objetivos se planteó la siguiente hipótesis general: “El sistema de análisis predictivo mejora la gestión de ventas en la empresa TAURO A&F SAC” y las siguientes hipótesis específicas:

HE1: El sistema de análisis predictivo incremento el índice de rotación de productos para la gestión de ventas de la empresa TAURO A&F SAC.

HE2: El sistema de análisis predictivo incremento el índice de ganancia de los productos para la gestión de ventas de la empresa TAURO A&F SAC.

HE3: El sistema de análisis predictivo incremento el índice de precisión de productos para la gestión de ventas de la empresa TAURO A&F SAC.

II. MARCO TEÓRICO

En la investigación realizada respecto al tema se dio un enfoque desde el ámbito internacional, nacional y local, considerándose las siguientes investigaciones:

En el plano **internacional** para GARCES (2020), en su investigación titulada “Aplicación de Machine Learning en la demanda de productos permisibles” en Guayaquil - Ecuador, donde se analiza la viabilidad que tiene la aplicación del Aprendizaje Automático en la previsión de la demanda de productos perecibles para lo cual se analizan las técnicas de algoritmos para que se haga una proyección de los factores comunes que apoye en las decisiones del ámbito productivo y comercial de la organización. Para la implementación de la solución se elaboró una red neuronal que pueda estudiar los patrones entre las variables identificadas y se pueda tomar la decisión más adecuada.

Suárez y Cárdenas (2017) en su investigación titulada “La rotación de inventarios y su incidencia en el flujo de efectivo” en Guayaquil - Ecuador, se analiza la fórmula para hallar la rotación de inventarios así como la máxima y mínima rotación de los productos para posteriormente realizar la valoración del stock e inducir en los costos que se deducen de ello.

Ángel y Márquez (2018) en su investigación titulada “Modelo de Gestión para la proyección de demanda de productos perecederos utilizando combinación de pronósticos por series de tiempo”, se analiza el problema de la organización respecto a los productos pendientes de rotación y en los costos que ocasiona. Desde el mes de septiembre del 2017 a septiembre del 2018 se produjo un incremento en los costos de \$681.400 a \$774.100 lo que equivale a un incremento del 13.6% en ese rango de fechas. El costo de mantener el inventario anualmente es \$9289.200 añadiendo el costo de ventas no realizadas por \$13529.724 mensuales que dejó de recibir la organización por la no venta de productos, para un total de \$171645.900 anual de pérdida. Este es un motivo para que se desarrolle un modelo que pueda incrementar el índice de rotación de productos que nos permita minimizar los gastos asociados al almacenaje.

Praveen (2020) en la investigación titulada “Inventory Management using Machine Learning”, se analiza el uso del modelo de regresión XGBoost (Extreme Gradient Boosting) para las predicciones de la demanda de productos, siendo este modelo un algoritmo de Aprendizaje Automático que usa árboles de decisión. Se decidió

hacer uso porque tenían información estructurada en la que usar árboles de decisión es lo ideal. Como resultado redujeron el stock y el capital destinado en las empresas.

Jaramillo (2020) en la investigación titulada “Gestión de pedidos para puntos de venta con aprendizaje automático”, se enfoca en la mecanización de las solicitudes teniendo como enfoque la demanda de productos para calcular la rotación del inventario en los puntos de venta y evitar el sobrestock. Se usaron los modelos estadísticos de Regresión Lineal, Holt-Winter junto con los modelos de Deep Learning Perceptron Multicapa y redes neuronales recurrentes. Se hizo uso de la metodología CRISP-DM siendo el modelo analítico más usado para proyectos de minería de datos.

Nacionales

Mamani (2019), en su investigación titulada “Modelo de pronóstico univariante de ventas mensuales de productos de la división de Electro hogar en Tiendas Peruanas S.A. Oechsle” en Huancayo, se analiza la forma de describir y predecir las ventas mensuales de los productos usando la metodología Box Jenkins para validar los modelos. Los datos fueron almacenados en base a las ventas mensuales posteriormente se identificó el modelo usando la autocorrelación y función de autocorrelación parcial. Para el análisis del modelo univariante de una serie temporal se utilizó la metodología ARIMA, el cual es un método cuantitativo moderno para la predicción e implementación en redes neuronales.

Ccoyccosi(2021) en la investigación titulada “Propuesta de un modelo de machine learning para el pronóstico de la demanda de prendas de vestir en la Corporación Brusko S.A.C”, el análisis se efectúa bajo la metodología CRISP-DM, regresión lineal y los algoritmos de aprendizaje supervisado, teniendo como objetivos pronosticar la demanda de pantalones para aprovechar las oportunidades de venta y la implementación de la previsión de la demanda por medio del uso de la regresión lineal del Aprendizaje Automático teniendo como objetivo minimizar las pérdidas en ventas por no tener disponibilidad de productos. Siendo el modelo propuesto implementado en tres etapas el tratamiento, la modelación para el pronóstico esperado y, el entrenamiento y despliegue del modelo.

Shimabukuro (2021) en la investigación titulada “Análisis de pronóstico de demanda para poder gestionar el inventario aplicado al sector salud”, este problema se

genera en la gestión de inventarios por lo tanto en la disponibilidad de los productos ocasionando inquietud en las demandas posteriores, teniendo un enorme gasto en la compra de insumos o carencia de los mismos. Se tiene como objetivo específico la creación de un aplicativo web que manifieste los productos a adquirir en el futuro para poder minimizar los riesgos descritos en un inicio. Este proyecto se desarrolló con la metodología Design Thinking, la base de datos con el proveedor Cloud AWS, la implementación de los algoritmos con Python y Anaconda, por último para mostrar los datos se utiliza Power BI.

Becerra y Villareal (2021) en la investigación titulada "Data Mining para el modelo predictivo de ventas y servicios de mantenimiento en un concesionario automotriz ligero", el análisis está enfocado en el área de post venta y en cómo identificar los clientes que están dispuestos a llevar su vehículo a mantenimiento para poder ofrecerles promociones. Se tiene como objetivos específicos desarrollar un modelo de red neuronal en la predicción de ventas y árboles para la clasificación de consumidores en el servicio post venta. Para la implementación del proyecto se usaron técnicas de minería de datos entre redes neuronales con el lenguaje Python y árboles de decisión con el lenguaje R. Entre las librerías usadas se encuentra Keras con la métrica MSE (Mean Square Error).

Variable independiente: Sistema de análisis predictivo. Según el artículo publicado sobre Tendencias de Business Intelligence en 2022 Comunicae newswire (2022) la predicción es un método en la cual se toma datos existentes para predecir lo que podría pasar en un futuro tomando en cuenta diferentes variables, esto permite obtener un resultado óptimo de lo que se quiere analizar. Además Rosero et al. (2022) define al análisis predictivo como el análisis de datos del pasado que sirve para determinar las probabilidades de una ocurrencia sobre un evento, sirve de guía para implementar soluciones basadas en la probabilidad.

En los enfoques conceptuales se consideran las siguientes definiciones Machine Learning o aprendizaje automático. Microsoft (2022) lo define como una especialidad de la inteligencia artificial(IA) enfocada en el desarrollo de sistemas analíticos para identificar patrones en los datos, estos patrones crean un modelo de datos que aprenden o mejoran el rendimiento de los procesos analíticos e intentan imitar el comportamiento del ser humano, La cual involucra tres enfoques

de aprendizaje.1-Aprendizaje automático supervisado, estos algoritmos son los más utilizados, este modelo es entrenado con distintos datos y etiquetas la cual aumenta su capacidad para realizar una predicción de lo que se desea obtener.2- Aprendizaje automático no supervisado, tiene un enfoque más libre sin etiquetas ni estructuras, el cual un ordenador puede reconocer procesos y patrones complejos sin la participación cercana del ser humano esto implica un entrenamiento basado en datos con un resultado no específico definido.3-Aprendizaje por refuerzo, en este tipo de aprendizaje existe la intervención humana el cual introduce un conjunto de acciones permitidas, reglas y estados finales, el algoritmo debe aprender por experiencia y recompensa. Además Bagnato (2022) define a Machine Learning como un sub campo de la inteligencia artificial que busca la creación de algoritmos que mejoren y adquieran experiencia de manera autónoma, esto implica que el software que se crea no necesariamente tiene reglas o instrucciones definidas, sino que éstos mejoran o aprenden de manera automática.

DATA MINING. Para Microsoft (2022) es el proceso que busca identificar información que aporte valor en las enormes bases de datos, para ello se usa el enfoque matemático que nos sirve para hallar patrones ocultos en los datos, los cuales no se pueden reconocer fácilmente debido a que es muy complejo o por la inmensa cantidad de datos recopilados. La minería de datos se puede aplicar en diferentes ámbitos, como: pronóstico, probabilidad, recomendaciones, riesgo, búsqueda de agrupación y secuencias. Además Vallejo, Guevara y Medina (2018) “sostienen que la minería de datos tiene como meta principal analizar los datos desde diferentes dimensiones, puntos de vista, ordenada y clasificada, con la finalidad de transformarla en información útil. También mencionan que se hace alusión a un grupo de algoritmos que facilitan el entendimiento de diferentes patrones útiles que se hallan encubiertos en las bases de datos.”

Aprendizaje supervisado, según Moarri (2022) menciona que es donde se entrena a la maquina a través de datos etiquetados para que pueda aprender y encontrar el patrón de reconocimiento.

Metodología CRISP-DM. Para IBM (2021), la metodología CRISP-DM, es una de las más aprovechadas en el desarrollo de proyectos con Data Mining, esta metodología comprende de 6 fases:1-Comprensión del negocio, en esta fase se

debe obtener toda la información requerida de la situación comercial actual y los objetivos comerciales de la organización.2-Comprensión de los datos, en esta etapa se hace un estudio profundo sobre los datos disponibles a través de tablas y gráficos, de forma que se pueda obtener la calidad de datos y descripción de los resultados.3-Preparación de los datos, en esta etapa se emplea más tiempo y esfuerzo debido a que se preparan y empaquetan los datos para la minería.4-Modelado, en esta fase se ejecuta múltiples modelos de acuerdo a los parámetros predeterminados se hacen ajustes a los parámetros o se retorna a la fase previa para poder manipular y obtener el modelo requerido.5-Evaluación, en esta fase se evalúan los resultados con los criterios detallados en la fase inicial.6-Despliegue, en esta fase nos centramos en utilizar los nuevos conocimientos para implementar mejoras en la organización. También García, Sánchez y Otegi (2021) define a la metodología CRISP-DM como una de las metodologías que tiene una amplia utilización en el campo de análisis de datos, ya que constituye las etapas de desarrollo de los proyectos relacionados al tratamiento de datos, su ventaja principal en relación a otras metodologías es que toma en cuenta las necesidades, problemas y objetivos del negocio, en la mayoría de veces forma parte de los procesos de machine learning.

El lenguaje en el cual se va desarrollar el sistema es Python, según Milanese (2022) “cuando hablamos de ciencia de datos, los lenguajes más habituales que se utilizan en el presente son Python y R. (...) el trabajo del científico de datos consiste en la manipulación y visualización de los datos así como el aprendizaje automático y profundo: para eso, Python proporciona tantas librerías que ayudan en cada tarea desde el inicio y fin de un proyecto enfocado en datos.”

Jupyter Notebooks, “es una aplicación web de código abierto, una interfaz para ejecutar programas que nos permite escribir y compartir código Python y documentos y podemos hacer todo eso en un navegador web.” (Milanese, 2022)
Scikit-learn, “es una librería de Python que nos ofrece distintos algoritmos para poder desarrollar las distintas fases del modelo predictivo (...) Resalta lo avanzado del desarrollo de esta librería cuando se requiere colocar un modelo en producción.” (Amat, 2020)

Pandas, según Martínez (2020) sostiene que integra el grupo de librerías fundamentales de Python para el desarrollo de modelos teniendo como principal valor la lectura y transformación de los distintos orígenes de datos en entradas de información.

Keras, según Veliz (2020) lo define como una interfaz de programación de alto nivel ya que se puede procesar la idea y pasarnos al resultado en un tiempo reducido, su aplicación es de gran éxito en modelos de aprendizaje profundo, los paquetes de estos modelos se ejecutan sin inconvenientes en unidades de procesamiento cpu o gpu, la utilización de estos llega a resumirse en pasos como: Lectura y preparación de los datos, definición del modelo, compilación del modelo y entrenamiento y evaluación del modelo.

Algoritmo Prophet, según Alonso-Cortés (2022) menciona que es un modelo en Python que se usa para predecir series temporales con una tendencia y estacionalidad muy marcada además que es fácil de entender e interpretar para poder dar valor a los datos.

Streamlit, según Correoso (2020) menciona que es una librería de Python que facilita la creación e intercambio de aplicaciones web enfocada en el aprendizaje automático.

III. METODOLOGÍA

3.1. Tipo y diseño de investigación

3.1.1 Tipo de investigación

La presente investigación es de tipo aplicada, la cual según Arispe (2020, p.62) se define que “se centra en reconocer a través del conocimiento científico, los medios (metodologías, tecnologías y protocolos) por los cuales se puede contribuir a solucionar una necesidad reconocida, práctica y específica”.

3.1.2 Diseño de investigación

El diseño de la investigación es pre-experimental, porque las variables van a ser manipuladas para ser colocadas en pruebas. En la que se tienen las siguientes características: Arispe (2020, p.67) “1) Manipulación de una o más variables (independiente), 2) Permite medir la variable dependiente, 3) existe un control del experimento, 4) Su alcance es explicativo y 5) Cuenta con validez interna”.

El tipo de investigación es cuantitativa según Mata (2019) cuando se elabora “con el propósito de responder a las preguntas de investigación planteadas y cumplir con los objetivos del estudio”. La forma como se validará la hipótesis es mediante el pretest y postest, siendo el pretest donde se evalúa la condición actual y el postest donde se válida la hipótesis.

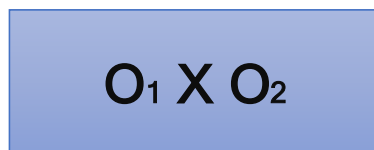

$$O_1 \times O_2$$

Figura 3: Formula para validar la hipótesis. Mata (2019)

O1: Gestión de ventas antes de la implementación del sistema de análisis predictivo (pretest).

X: Aplicación de sistema de análisis predictivo

O2: Gestión de ventas en la demanda de productos después de la implementación del sistema de análisis predictivo (postest)

3.2. Variables y operacionalización

La variable independiente es “Sistema de análisis predictivo” y la dependiente es “gestión de ventas”, teniendo está 3 dimensiones las cuales son: rotación de productos con el indicador índice de rotación de productos; rentabilidad de

productos con el indicador índice de ganancia de productos y precisión de ventas con el indicador índice de precisión de productos, los cuales se muestra en el **Anexo 1**. El detalle es el siguiente:

a) **Definición conceptual:** según Zambrano et al (2021) “La Gestión de Ventas tiene como objetivo principal mantener y aumentar las ventas mediante la gestión del talento humano en otras palabras traducir los objetivos y estrategias de la empresa en objetivos y estrategias de recursos, lo cual se logra mediante la planeación estratégica de ventas además del mercado potencial de clientes”.

b) **Definición operacional:** Permite reconocer y predecir los productos que más se venden.

c) **Dimensiones:** rotación de productos (Alvarez y Parada 2020, p.57), rentabilidad de productos (Pacheco 2019) y precisión de ventas (Villanueva 2021).

d) **Indicadores:** índice de rotación de productos (Alvarez y Parada 2020, p.57), índice de rentabilidad de productos (Pacheco 2019) e índice de precisión de productos (Villanueva 2021).

e) **Escala de medición:** razón.

3.3. Población, muestra y muestreo

3.3.1. Población

La población que abarca la presente investigación está integrada por los 212 registros de venta del producto que se vendió entre los años 2021-2022.

Criterio de inclusión: el estrato a analizar es el producto que tiene mayor cantidad de ventas

3.3.1. Muestra y Muestreo

Muestra

Para el tema investigado se usó toda la población.

Muestreo

Se empleó el muestreo probabilístico estratificado, esta muestra se segmenta por 30 días de muestra estratificada.

Según (Hernández y Carpio, 2019) la muestra probabilística estratificada quiere decir que “en este método se divide a la población en estratos o subgrupos menores, parecidos internamente respecto a una característica, pero heterogéneos entre ellos, diferenciándolos por una variable que resulte de interés para la investigación, por ejemplo, según la profesión, municipio, estado civil, sexo.”

3.4. Técnicas e instrumentos de recolección de datos

Técnica de recolección de datos

Según Hernández Mendoza y Duana Ávila (2020) son todas las actividades y procesos que permiten al investigador obtener la información indispensable para satisfacer las metas de la investigación.

Técnica: fichaje

Se utilizó para realizar la recopilación de la información de los indicadores.

Instrumentó de recolección de datos: fichas de Registro

Se emplearon formatos de fichas de registro para cada indicador, utilizando las formulas planteadas en el capítulo anterior.

Validez, representa el grado de confiabilidad del instrumento a usarse para la recopilación de datos. Los indicadores fueron validados en base a autores, siendo los siguientes:

- a) índice de rotación de productos (RP), Alvarez y Parada, 2020, p 57.
- b) índice de ganancia de los productos (U), Vallejos & Chilinguina, p. 57, 2017, citado por Pacheco 2019
- c) índice de precisión de productos (PVT), Villanueva, 2021.

Confiabilidad, según Villasis Miguel, et al. (2018) es la particularidad que se le atribuye al instrumento cuando no se tienen sesgos en la información recopilada.

En esta investigación, se usó el método de confiabilidad utilizando el software IBM SPSS Statistics 27.

La validación de los instrumentos por expertos se encuentra en el **Anexo 11**

3.5. Procedimientos

El software que gestiona las transacciones de las ventas se encuentra en Microsoft SQL Server el cual administra el proveedor del aplicativo, para obtener estos datos el proveedor envía reportes mensuales de las ventas en un archivo Excel.

Para poder recopilar esta información se utilizó la técnica del fichaje para posteriormente generar las fichas de registro por indicador.

3.6. Método de análisis de datos

Para el análisis de los datos tanto en el pretest y postest se usó como apoyo el aplicativo SPSS v27.

Respecto al comportamiento de las variables en el ámbito descriptivo se usará tablas y figuras que muestren los datos relacionados al instrumento.

Respecto al comportamiento de las variables en el ámbito inferencial se empleó prueba de rangos con signo de Wilcoxon en lo relacionado a las pruebas no paramétricas.

3.7. Aspectos éticos

La presente investigación desarrollada por los autores tiene información veraz y se verifica para la implementación de la solución en la organización además nos basamos en el código de ética del colegio de ingenieros del Perú.

IV. RESULTADOS

Análisis descriptivo

Luego que se observaron los resultados recopilados en las fases de pre test y post test para obtener las medidas descriptivas de los indicadores: índice de rotación de productos, índice de ganancia de productos e índice de precisión de productos para antes y después de implementar aprendizaje automático.

Tabla 1: Medidas de los indicadores

Indicador	N	Mínimo	Máximo	Media	Desviación estándar
Índice de rotación de productos - Pretest	30	0	0.09	0.01	0.03
Índice de rotación de productos - Postest	30	0	0.06	0.03	0.02
Índice de ganancia de productos - Pretest	30	0	82.5	9.9	22.35
Índice de ganancia de productos - Postest	30	0	49.5	24.75	13.53
Índice de precisión de productos - Pretest	30	-1	0.80	0.05	0.30
Índice de precisión de productos - Postest	30	0	1.00	0.14	0.25

Fuente: Elaboración propia en base al aplicativo estadístico IBM SPSS

Indicador 1: Índice de rotación de productos

En la siguiente imagen, se presenta la conducta de la variable, antes y después de implementar el aprendizaje automático, se puede inferir que el índice de rotación de productos mejoro de un 1% a 3% por lo cual la mejora fue de un 2%.

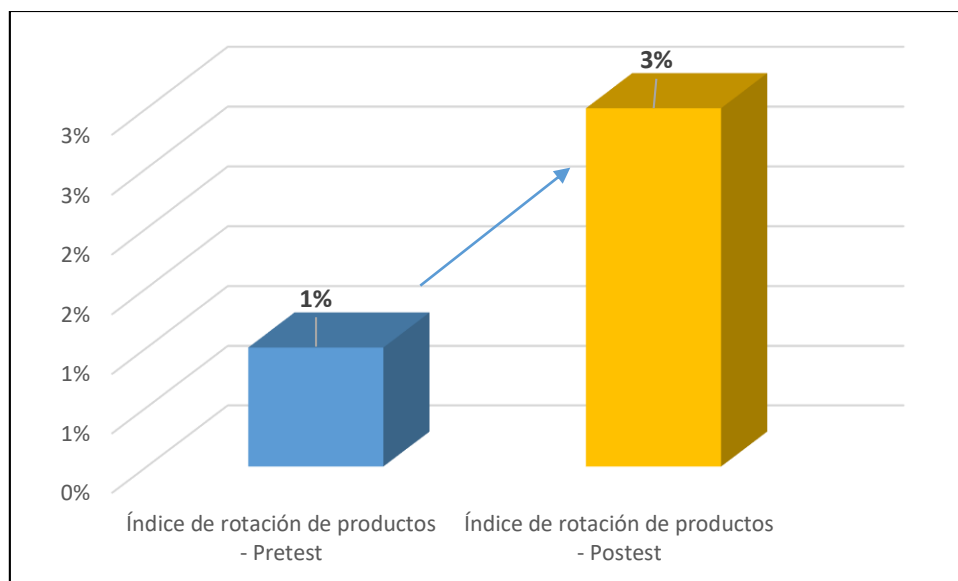


Figura 4: Pretest y postest del índice de rotación de productos (Elaboración propia)

Indicador 2: Índice de ganancia de productos

En la siguiente imagen, se presenta la conducta de la variable, antes y después de implementar el aprendizaje automático, se puede inferir que el índice de ganancia de productos mejoro de un 10% a 25% por lo cual la mejora fue de un 15%.

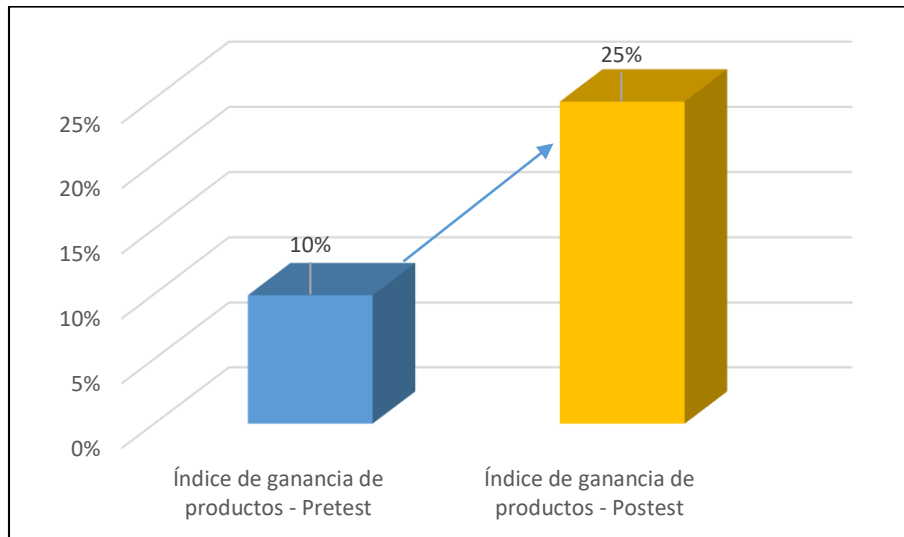


Figura 5: Pretest y postest del índice de ganancia de productos (Elaboración propia)

Indicador 3: Índice de precisión de productos

En la siguiente imagen, se presenta la conducta de la variable, antes y después de implementar el aprendizaje automático, se puede inferir que el índice de precisión de productos mejoro de un 5% a 14% por lo cual la mejora fue de un 9%.

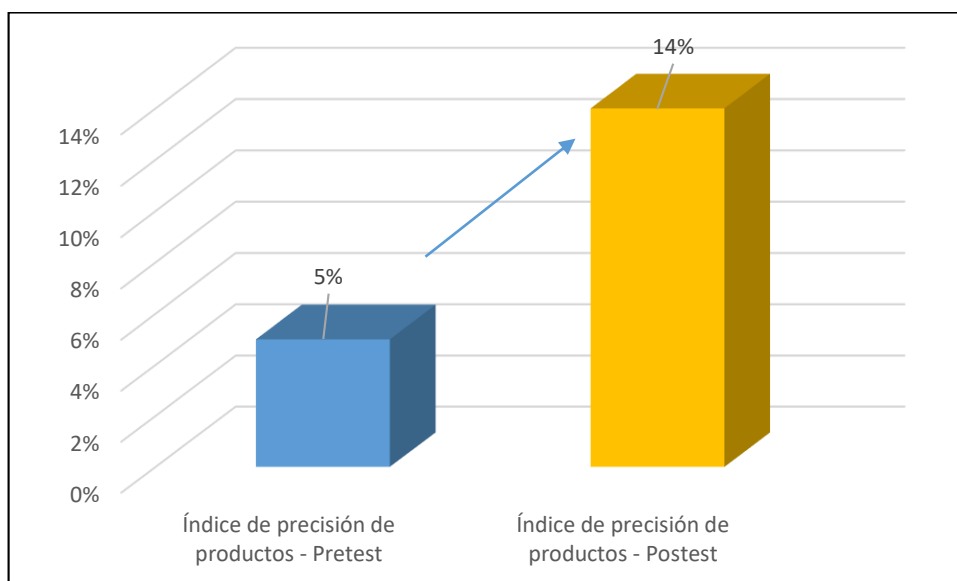


Figura 6: Pretest y postest del índice de precisión de productos (Elaboración propia)

Análisis inferencial

Formulación de la hipótesis estadística específica (HE1) 1:

HEo: El sistema de análisis predictivo no incrementará el índice de rotación de productos para la gestión de ventas de la empresa TAURO A&F SAC.

HE1: El sistema de análisis predictivo incrementará el índice de rotación de productos para la gestión de ventas de la empresa TAURO A&F SAC.

El análisis se realizó a 212 registros de venta segmentándose para 30 días para obtener detalles descriptivos de la variable índice de rotación de productos se encuentra en el apartado “Análisis Descriptivo”.

Prueba de normalidad

Se realizó con el método de Shapiro-Wilk porque se analizó en 30 registros.

Tabla 2: Prueba de normalidad del índice de rotación de productos

Indicador	Shapiro-Wilk		
	Estadístico	gl	Sig.
Índice de rotación de productos antes	0.51	30	0
Índice de rotación de productos después	0.84	30	0

Fuente: Elaboración propia en base al aplicativo estadístico IBM SPSS

Donde:

Índice de rotación de productos antes:

En la tabla 2, se presenta la prueba de normalidad dando el resultado del pretest donde el nivel de significancia es igual a 0.00 lo cual comprueba que no coincide con la distribución normal.

Índice de rotación de productos después:

En la tabla 2, se presenta la prueba de normalidad dando el resultado del postest donde el nivel de significancia es igual a 0.00 lo cual comprueba que no coincide con la distribución normal.

Prueba de rangos con signos de Wilcoxon

En la siguiente tabla, se presenta la aplicación de la prueba Wilcoxon en el índice de rotación de productos.

Tabla 3: Prueba Wilcoxon del indicador índice de rotación de productos

		N	Rango promedio	Suma de rangos
Índice de rotación de productos después - Índice de rotación de productos antes	Rangos negativos	5 ^a	16.60	83.00
	Rangos positivos	23 ^b	14.04	323.00
	Empates	2 ^c		
	Total	30		

Fuente: Elaboración propia en base al aplicativo estadístico IBM SPSS

- a. Índice de rotación de productos después < Índice de rotación de productos antes
- b. Índice de rotación de productos después > Índice de rotación de productos antes
- c. Índice de rotación de productos después = Índice de rotación de productos antes

En la siguiente tabla, se presentan los resultados con la prueba Z en el índice de rotación de productos.

Tabla 4: Prueba Z del indicador índice de rotación de productos

	Índice de rotación de productos después - Índice de rotación de productos antes
Z	-2,792 ^b
Sig. asin. (bilateral)	0.005

Fuente: Elaboración propia en base al aplicativo estadístico IBM SPSS

- a. Prueba de rangos con signo de Wilcoxon
- b. Se basa en rangos positivos

Después de analizar los datos con el software SPSS, podemos verificar en la tabla 4 que se tiene el valor de Z igual a -2.792, por lo que se encuentra en el área de rechazo donde se obtiene el valor de $p=0.005 < 0.05$, se infiere que se rechazó la H₀ y se aceptó la H₁; por lo tanto, se aceptó que “El sistema de análisis predictivo incrementará el índice de rotación de productos para la gestión de ventas de la empresa TAURO A&F SAC” con un incremento de 2%.

Formulación de la hipótesis estadística específica (HE2) 2:

HEo: El sistema de análisis predictivo no incrementará el índice de ganancia de los productos para la gestión de ventas de la empresa TAURO A&F SAC.

HE1: El sistema de análisis predictivo incrementará el índice de ganancia de los productos para la gestión de ventas de la empresa TAURO A&F SAC

El análisis se realizó a 212 registros de venta segmentándose para 30 días para obtener detalles descriptivos de la variable índice de ganancia de productos se encuentra en el apartado “Análisis Descriptivo”.

Prueba de normalidad

Se realizó con el método de Shapiro-Wilk porque se analizó en 30 registros.

Tabla 5: Prueba de normalidad del índice de ganancia de productos

Indicador	Shapiro-Wilk		
	Estadístico	gl	Sig.
Índice de ganancia de productos antes	0.51	30	0
Índice de ganancia de productos después	0.84	30	0

Fuente: Elaboración propia en base al aplicativo estadístico IBM SPSS

Donde:

Índice de ganancia de productos antes:

En la tabla 5, se presenta la prueba de normalidad dando el resultado del pretest donde el nivel de significancia es igual a 0.00 lo cual comprueba que no coincide con la distribución normal.

Índice de ganancia de productos después:

En la tabla 5, se presenta la prueba de normalidad dando el resultado del postest donde el nivel de significancia es igual a 0.00 lo cual comprueba que no coincide con la distribución normal.

Prueba de rangos con signos de Wilcoxon

En la siguiente tabla, se presenta la aplicación de la prueba Wilcoxon en el índice de ganancia de productos.

Tabla 6: Prueba Wilcoxon del índice de ganancia de productos

		N	Rango promedio	Suma de rangos
Índice de ganancia de productos después - Índice de ganancia de productos antes	Rangos negativos	5 ^a	16.60	83.00
	Rangos positivos	23 ^b	14.04	323.00
	Empates	2 ^c		
	Total	30		

Fuente: Elaboración propia en base al aplicativo estadístico IBM SPSS

- a. Índice de ganancia de productos después < Índice de ganancia de productos antes
- b. Índice de ganancia de productos después > Índice de ganancia de productos antes
- c. Índice de ganancia de productos después = Índice de ganancia de productos antes

En la siguiente tabla, se presentan los resultados con la prueba Z en el índice de ganancia de productos.

Tabla 7: Prueba Z del índice de ganancia de productos

	Índice de ganancia de productos después - Índice de ganancia de productos antes
Z	-2,792 ^b
Sig. asin. (bilateral)	0.005

Fuente: Elaboración propia en base al aplicativo estadístico IBM SPSS

- a. Prueba de rangos con signo de Wilcoxon
- b. Se basa en rangos positivos

Después de analizar los datos con el software SPSS, podemos verificar en la tabla 7 que se tiene el valor de Z igual a -2.792, por lo que se encuentra en el área de rechazo donde se obtiene el valor de $p=0.005 < 0.05$, se infiere que se rechazó la H₀ y se aceptó la H₁; por lo tanto, se aceptó que “El sistema de análisis predictivo incrementará el índice de ganancia de los productos para la gestión de ventas de la empresa TAURO A&F SAC” con un incremento de 15%.

Formulación de la hipótesis estadística específica (HE3) 3:

HE₀: El sistema de análisis predictivo no incrementará el índice de precisión de productos para la gestión de ventas de la empresa TAURO A&F SAC.

HE₁: El sistema de análisis predictivo incrementará el índice de precisión de productos para la gestión de ventas de la empresa TAURO A&F SAC.

El análisis se realizó a 212 registros de venta segmentándose para 30 días para obtener detalles descriptivos de la variable índice de precisión de productos se encuentra en el apartado “Análisis Descriptivo”.

Prueba de normalidad

Se realizó con el método de Shapiro-Wilk porque se analizó en 30 registros.

Tabla 8: Prueba de normalidad del índice de precisión de productos

Indicador	Shapiro-Wilk		
	Estadístico	gl	Sig.
Índice de precisión de productos antes	0.55	30	0.00
Índice de precisión de productos después	0.63	30	0.00

Fuente: Elaboración propia en base al aplicativo estadístico IBM SPSS

Dónde:

Índice de precisión de productos antes:

En la tabla 8 se presenta la prueba de normalidad dando el resultado del pretest donde el nivel de significancia es igual a 0.00 lo cual comprueba que no coincide con la distribución normal.

Índice de precisión de productos después:

En la tabla 8 se presenta la prueba de normalidad dando el resultado del postest donde el nivel de significancia es igual a 0.00 lo cual comprueba que no coincide con la distribución normal.

Prueba de rangos con signos de Wilcoxon

En la siguiente tabla, se presenta la aplicación de la prueba Wilcoxon en el índice de precisión de venta.

Tabla 9: Prueba Wilcoxon del índice de precisión de venta

		N	Rango promedio	Suma de rangos
Índice de precisión de productos después - Índice de precisión de productos antes	Rangos negativos	0 ^a	0.00	0.00
	Rangos positivos	30 ^b	15.50	465.00
	Empates	0 ^c		
	Total	30		

Fuente: Elaboración propia en base al aplicativo estadístico IBM SPSS

a. Índice de precisión de productos después < Índice de precisión de productos

antes

b. Índice de precisión de productos después > Índice de precisión de productos antes

c. Índice de precisión de productos después = Índice de precisión de productos antes

En la siguiente tabla, se presentan los resultados con la prueba Z en el índice precisión de productos.

Tabla 10: Prueba Z del índice de precisión de venta

	Índice de precisión de productos después - Índice de precisión de productos antes
Z	-4,782 ^b
Sig. asin. (bilateral)	0.000

Fuente: Elaboración propia en base al aplicativo estadístico IBM SPSS

a. Prueba de rangos con signo de Wilcoxon

b. Se basa en rangos positivos

Después de analizar los datos con el software SPSS, podemos verificar en la tabla 10 que se tiene el valor de Z igual a -4.782, por lo que se encuentra en el área de rechazo y se obtiene el valor de $p=0.000 < 0.05$, se infiere que se rechazó la H₀ y se aceptó la H₁; por lo tanto, se aceptó que “El sistema de análisis predictivo incrementará el índice de precisión de productos para la gestión de ventas de la empresa TAURO A&F SAC” con un incremento de 9%.

Prueba de hipótesis general

Luego de validar anteriormente que se aceptó las hipótesis de los objetivos específicos planteados en un inicio; por lo tanto, se acepta la hipótesis general: “El sistema de análisis predictivo mejora la gestión de ventas en la empresa TAURO A&F SAC”.

Resumen

ID	HIPÓTESIS	RESULTADO
HE1	El sistema de análisis predictivo incrementará el índice de rotación de productos para la gestión de ventas de la empresa TAURO A&F SAC	Aceptado

HE2	El sistema de análisis predictivo incrementará el índice de ganancia de los productos para la gestión de ventas de la empresa TAURO A&F SAC	Aceptado
HE3	El sistema de análisis predictivo incrementará el índice de precisión de productos para la gestión de ventas de la empresa TAURO A&F SAC	Aceptado

V. DISCUSIÓN

La actual investigación en la validación de la primera hipótesis presento los siguientes hallazgos, obtuvo como resultados que el sistema de análisis predictivo con aprendizaje automático aumento el índice de rotación de productos de 1% a 3% teniendo una diferencia en 2%. En la tabla 1 se presentan los valores descriptivos del indicador en mención.

Del mismo modo Bernaola y Varillas (2022), en su investigación titulada “Sistema predictivo con Machine Learning para la gestión de inventario para la Empresa Inversiones Ferreteras Mendoza” menciona que el indicador de rotación de inventario se incrementó de 0.87% a 0.93% incrementándose en un 0.06%.

Además Silva (2019), en su investigación titulada “Aplicación Web Para El Control de Inventario En La Empresa Popeyito Del Distrito De Ate” menciona que el indicador rotación de inventario se incrementó de 0.85 a 1.6 incrementándose en 0.75.

En la validación de la segunda hipótesis presento los siguientes hallazgos, obtuvo como resultados que el sistema de análisis predictivo con aprendizaje automático aumento las ganancias de productos de 10% a 25% teniendo una diferencia de 15%. En la tabla 1 se presentan los valores descriptivos del indicador en mención. Del mismo modo Coronel y Cáceres (2019), en su investigación titulada “Automatización del Sistema de Información aplicando Machine Learning en el área logística en Villa Chicken S.A.C” menciona que el indicador crecimiento de ventas se incrementó de 32.50% a 46.71% incrementándose en un 14.21%, concluyendo que el aprendizaje automático influyo en el crecimiento del indicador.

En la validación de la tercera hipótesis presento los siguientes hallazgos, obtuvo como resultados que el sistema de análisis predictivo con aprendizaje automático aumento la precisión de productos de 5% a 14% teniendo una diferencia de 9%. En la tabla 1 se presentan los valores descriptivos del indicador en mención.

Del mismo modo Romero (2021), en su investigación titulada “Inteligencia de negocio en el proceso de ventas de la empresa Q System S.A.C” menciona que el indicador crecimiento de ventas se incrementó de 5,26% a 114.39% incrementándose en un 109.13%, concluyendo que la inteligencia de negocio influyo en el crecimiento del indicador.

En cuanto al objetivo general de lo investigado se busca determinar la influencia del sistema de análisis predictivo para la gestión de ventas de la empresa TAURO A&F SAC, podemos determinar que hubo un incremento de 2% en el indicador de rotación y 15% en la ganancia de productos además en el indicador precisión de productos se incrementó en 9%.

VI. CONCLUSIONES

Luego de lo investigado se obtuvieron las siguientes conclusiones:

- 1) Se concluye que posteriormente a la implementación del aprendizaje automático se mejora la gestión de ventas expresándose por medio de los indicadores índice de rotación de productos, índice de ganancia de productos y precisión de venta, que fue el objetivo de la investigación.
- 2) Respecto al primer indicador índice de rotación de productos, se obtuvo un incremento de 2%, debido a que se incluyó en el sistema la cantidad de productos a vender en determinada fecha.
- 3) Respecto al segundo indicador índice de ganancia de productos, se obtuvo un incremento de 15% debido al incremento en la rotación de los mismos.
- 4) Respecto al tercer indicador índice de precisión de productos, se obtuvo un incremento de 9% debido a la predicción del producto a vender.

VII. RECOMENDACIONES

Luego de lo investigado se recomienda lo siguiente:

- 1) Se tiene que tener acceso libre a la base de datos para poder hacer conexiones directas y poder recopilar la información de manera eficiente.
- 2) Para realizar el análisis de ventas de productos se tiene que tener un amplio historial de ventas de aproximadamente 3 años para que el modelo de machine learning pueda procesar mejor la información y así disminuir el sesgo.
- 3) El proceso de ventas debe estar completamente integrado con el sistema de inventario para poder tener mejor control del ingreso y salida de productos para poder así recopilar la información y generar reportes que apoyen en la en las decisiones de la organización.

REFERENCIAS

ALONSO-CORTES, Marina, 2022. Análisis y Predicción de Series Temporales con FB Prophet Python. En: *modeldifferently* [en línea]. Disponible en: https://www.modeldifferently.com/2022/04/analisis_prediccion_ts_prophet/

ÁLVAREZ, Luis y PARADA, Sandra, 2020. *Gestión de inventarios: cartilla para el aula* [en línea]. Bogota: Corporación Universitaria Minuto de Dios [consulta: 25 de abril de 2022]. ISBN 9789587634341. Disponible en: <https://repository.uniminuto.edu/handle/10656/11481>

AMAT, Joaquín, 2020. *Machine learning con Python y Scikit-learn* [en línea]. Disponible en: https://www.cienciadedatos.net/documentos/py06_machine_learning_python_scikitlearn.html

ANGEL, Nelson y MÁRQUEZ, Diana, 2018. *Modelo de gestión para la proyección de demanda de productos perecederos utilizando combinación de pronósticos por series de tiempo* [en línea]. Tesis. Bogotá: Universidad Sergio Arboleda [consulta: mayo de 2022]. Disponible en: <https://repository.usergioarboleda.edu.co/handle/11232/1805>

ARISPE ALBURQUEQUE, Claudia M., et al., 2020. *La Investigación Científica. Una aproximación para los estudios de posgrado* [en línea]. Guayaquil: Universidad Internacional del Ecuador [consulta: 5 mayo 2022]. ISBN 9789942385789. Disponible en: <https://repositorio.uide.edu.ec/bitstream/37000/4310/1/LA%20INVESTIGACION%20CIENTIFICA.pdf>

BAGNATO, Juan, 2022. *Aprende Machine Learning*. [en línea]. España: [consulta: 01 mayo 2022]. ISBN: 9788409258161. Disponible en: <https://leanpub.com/aprendeml>

BECERRA, Jocelyn y VILLARREAL, Enrique, 2021. *Data Mining para modelo*

predictivo de ventas y servicios de mantenimiento en un concesionario automotriz ligero [en línea]. Tesis. Lima: Universidad de Lima [consulta: mayo de 2022]. Disponible en: <https://repositorio.ulima.edu.pe/handle/20.500.12724/15395>

BERNAOLA, Danny y VARILLAS, Pablo, 2022. *Sistema predictivo con Machine Learning para la gestión de inventario para la Empresa Inversiones Ferreteras Mendoza S.A.C* [en línea]. Tesis. Universidad Cesar Vallejo [consulta: 15 de noviembre del 2022]. Disponible en: <https://hdl.handle.net/20.500.12692/97798>

CASTIGLIONE, Raúl y LAZARO, Jesús, 2019. Sistema web para el proceso de control de almacén de la empresa Tic Integrity G & V S.A.C. [en línea]. Tesis. Universidad Cesar Vallejo [consulta: 15 de noviembre del 2022]. Disponible en: <https://hdl.handle.net/20.500.12692/41949>

CCOYCCOSI, Ronald A., et al., 2021. *Propuesta de un modelo de Machine Learning para el pronóstico de la demanda de prendas de vestir en la Corporación Brusko S.A.C* [en línea]. Tesis. Lima: Universidad Esan [consulta: abril de 2022]. Disponible en: <https://hdl.handle.net/20.500.12640/2931>

CHÁVEZ, Elsita, et al., 2018. Inteligencia Artificial en la toma de decisiones gerenciales. (Spanish). *Dilemas Contemporáneos: Educación, Política y Valores* 1–12. [consulta: mayo 2022]. Disponible en: <http://search.ebscohost.com/login.aspx?direct=true&db=eue&AN=130757964&lang=es&site=eds-live>

COMUNICAE NEWSWIRE, 2022. Tendencias de Business Intelligence en 2022 según DigDash, el software de visualización y análisis de datos [en línea]. En: *ProQuest*. Disponible en: <https://www.proquest.com/docview/2621224670?accountid=37408&forcedol=true> [consulta: 30 de junio de 2022]

CORONEL, Yoel y CACERES, Walther, 2019. *Automatización del Sistema de Información aplicando Machine Learning en el área logística en Villa Chicken S.A.C*

[en línea]. Tesis. Universidad Cesar Vallejo [consulta: 15 de noviembre del 2022]. Disponible en: <https://hdl.handle.net/20.500.12692/47899>

ESCUADERO, Francisco. Tercera edición del estudio sobre la percepción de madurez digital en las empresas del Perú. *Ernst & Young Global Ltd* [en línea]. 27 de abril del 2021. Disponible en: https://www.ey.com/es_pe/consulting/madurez-digital-en-peru [consulta: 15 de Setiembre de 2021].

FIGUEROA, Renzo, 2019. *Sistema de Información utilizando Machine Learning para la Predicción Financiera en la Empresa Cosmos SAC Lima - 2019* [en línea]. Tesis. Universidad Cesar Vallejo [consulta: 15 de noviembre del 2022]. Disponible en: <https://hdl.handle.net/20.500.12692/70064>

GARCES, Hans, 2020. *Aplicación de Machine Learning en la demanda de productos permisibles* [en línea]. Tesis titulación grado. Ecuador: Universidad Católica de Santiago de Guayaquil [consulta: 15 de Setiembre de 2021]. Disponible en: <http://repositorio.ucsg.edu.ec/handle/3317/14264>

GARCÍA, Imanol, SÁNCHEZ, Francisco y OTEGI, Jose, 2021. Metodología para la definición de tendencias de aplicación en tecnologías emergentes: Blockchain como caso de estudio. *Revista Española De Documentación Científica* [en línea]. CSIC, Vol. 44 no.4. Pp 1-19 [Fecha consulta: 27 abril 2022]. ISSN: 0210-0614 Disponible en: <https://doi.org/10.3989/redc.2021.4.1771>

HERNÁNDEZ, Carlos y CARPIO, Natalia, 2019. Introducción a los tipos de muestreo. *Alerta, Revista científica Del Instituto Nacional De Salud* [en línea]. San Salvador: Vol.2, nº1. pp 76 - 79 [consulta: 10 mayo 2022]. ISSN 2617-5274. Disponible en: <https://doi.org/10.5377/alerta.v2i1.7535>

HERNANDEZ, Sandra y DUANA, Danae, 2020. *Técnicas e instrumentos de recolección de datos. Boletín Científico de las Ciencias Económico Administrativas del ICEA.* [en línea] Vol.9, No. 17,1-3[consulta: 20 Agosto 2022] ISSN: 2007-4913. Disponible en:

<https://repository.uaeh.edu.mx/revistas/index.php/icea/article/download/6019/7678>

HIGA, Hugo, 2021. *Análisis de pronóstico de demanda para poder gestionar el inventario aplicado al sector salud* [en línea]. Tesis. Lima: Universidad de Lima [consulta: mayo de 2022]. Disponible en: <https://hdl.handle.net/20.500.12724/14099>

HUNWITS, Judith y KRISH, Daniel, 2018. *Machine Learning for dummines* [en línea]. [Consulta 20 de abril del 2022]. ISBN 978-1-119-45495-3. Disponible en: <https://www.ibm.com/downloads/cas/GB8ZMQZ3>

IBM, 2021. *Introducción al CRISP-DM* [en línea]. En: *IBM*. Disponible en: <https://www.ibm.com/docs/es/spss-modeler/saas?topic=guide-introduction-crisp-dm> [consulta: 17 mayo 2022].

JARAMILLO, Juan, 2020. *Gestión de pedidos para puntos de venta con aprendizaje Automático* [en línea]. Tesis. Bogotá: Universidad de los Andes [consulta: abril de 2022]. Disponible en: <http://hdl.handle.net/1992/49180>

LOHR, Steve, 2018. *La promesa de la inteligencia artificial para transformar industrias*. The New York Times [en línea]. 22 de octubre del 2018. Disponible en: <https://www.nytimes.com/es/2018/10/22/espanol/inteligencia-artificial-industrias.html> [consulta: 15 de Setiembre de 2021].

LOPEZ, Daniela, et al., 2021. *Comparación de arquitecturas de redes neuronales convolucionales para el diagnóstico de COVID-19*. *Scielo* [en línea]. Lugar: Comp. y Sist, 25, 3 [Fecha consulta: 27 Mayo 2022]. ISSN 2007-9737. Disponible en: https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-55462021000300601#aff1

MAMANI, Basilio, 2019. *Modelo de pronostico univariante de ventas mensuales de productos de la división de electro-hogar en tiendas Peruanas S.A. Oechsle* [en línea]. Tesis. Puno: Universidad Nacional del Altiplano [consulta: mayo de 2022]. Disponible en: <http://repositorio.unap.edu.pe/handle/20.500.14082/13082>

MANSANET, Jordi y PÉREZ Jaime, 2020. Librerías imprescindibles de Python para Machine Learning. En: *Solver Intelligent Analytics* [en línea]. Disponible en: <https://iasolver.es/6-librerias-de-python-para-machine-learning/> [consulta: 18 junio 2022].

MATA, Luis, 2019. Los diseños de investigaciones con enfoque cuantitativo. En: *Investigalia* [en línea]. Disponible en: <https://investigaliacr.com/investigacion/los-disenos-de-investigaciones-con-enfoque-cuantitativo/> [consulta: 10 Octubre 2022]

MICROSOFT, 2022. Conceptos de minería de datos. En: *Microsoft* [en línea]. Disponible en: <https://learn.microsoft.com/es-es/analysis-services/data-mining/data-mining-concepts?view=asallproducts-allversions> [consulta: 21 de mayo de 2022]

MILANESE, Stefano. 2022. Python para la ciencia de datos. En: MILANESE, Stefano. *Ciencia de datos y aprendizaje automático: Técnicas y códigos en Python comentados*. Lugar de publicación: Ed. Español, 63 ASIN B09XTS5L52

MOARRI, Richard, 2022. Machine learning: Cuáles son sus tipos y en qué se diferencian. En: *ConexiónEsan* [en línea]. Disponible en: <https://www.esan.edu.pe/conexion-esan/machine-learning-cuales-son-sus-tipos-y-en-que-se-diferencian> [consulta: 19 Mayo 2022]

PACHECO, Fabiola, 2019. *Módulo costos de producción* [en línea]. Universidad Santo Tomás Seccional Tunja: Ediciones Usta [consulta: 01 mayo 2022]. ISBN 9789585471269. Disponible en: <http://hdl.handle.net/11634/18470>

PRAMOD, Singh y MANURE, Avinash, 2020. *Learn TensorFlow 2.0: Implement Machine Learning and Deep Learning Models with Python* [en línea]. 1 ed. Bangalore: Apress. [Consulta: abril 2022]. ISBN 978-1-4842-5558-2. Disponible en: <https://link.springer.com/book/10.1007/978-1-4842-5558-2>.

PRAVEEN K, B et al., 2020. Inventory Management using Machine Learning. *International Journal of Engineering Research and Technology (IJERT)* [en línea]. Gandhinagar: Editorial Management System (EMS), vol. 9 no. 6, Pp 866 – 869 [consulta: 15 de abril de 2022]. ISSN 2278-0181. Disponible en: <https://www.ijert.org/research/inventory-management-using-machine-learning-IJERTV9IS060661.pdf>

ROSERO, Paul, et al., 2017. Arquitectura, servicios y aplicaciones de Business Intelligence Revisión de literatura en: *ResearchGate* [en línea]. Disponible en: https://www.researchgate.net/publication/313280378_ARQUITECTURA_SERVICIOS_Y_APLICACIONES_DE_Business_Intelligence_Revisión_de_literatura

SILVA, Sergio, 2019. *Aplicación Web Para El Control de Inventario En La Empresa Popeyito Del Distrito De Ate* [en línea]. Tesis. Universidad Cesar Vallejo [consulta: 15 de noviembre del 2022]. Disponible en: <https://hdl.handle.net/20.500.12692/86570>

SUÁREZ, Gastón y CÁRDENAS Polet. 2017. La rotación de los inventarios y su incidencia en el flujo de efectivo. *Revista Observatorio de la Economía Latinoamericana* [en línea]. Ecuador: Universidad Laica Vicente Rocafuerte de Guayaquil, pp. 1-12 [consulta: 5 de mayo de 2022]. ISSN: 1696-8352. Disponible en: <http://www.eumed.net/cursecon/ecolat/ec/2017/inventarios-flujo-efectivo.html>

TORRES, Jordi, 2019. Redes Neuronales Recurrentes En: *Supercomputing and artificial intelligence* [en línea]. Disponible en: <https://torres.ai/redes-neuronales-recurrentes/> [consulta: 20 de mayo de 2022]

VALLEJO, Henry, GUEVARA, Edelmira y MEDINA, Segundo, 2018. Minería de Datos. *RECIMUNDO: Revista Científica de la Investigación y el Conocimiento* [en línea]. Honduras: Editorial Saberes del Conocimiento, vol.2 no. pp 339-349 [consulta: 28 abril 2022]. ISSN 2588-073X. Disponible en: <https://www.recimundo.com/index.php/es/article/view/182>

VÉLIZ, Carlos, 2020. *Aprendizaje automático Introducción al aprendizaje profundo* [en línea]. Lima: El Fondo Editorial de la Pontificia Universidad Católica del Perú [consulta: 3 junio 2022]. ISBN: 9786123175887 Disponible en: https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=2600876&lang=es&site=eds-live&ebv=EB&ppid=pp_262

VILLASIS, Miguel, et al., 2018. El protocolo de la investigación VII. Validez y confiabilidad de las mediciones. *Revista Alergia México* [en línea]. Lugar: 7 páginas [consulta: 20 de Octubre del 2022]. Disponible en: <https://www.scielo.org.mx/pdf/ram/v65n4/2448-9190-ram-65-04-414.pdf>

ZAFORAS, Manuel, 2017. Tensorflow, o cómo será el futuro de la inteligencia artificial según google. En *Paradigma* [en línea]. Disponible en: <https://www.paradigmadigital.com/dev/tensorflow-sera-futuro-la-inteligencia-artificial-segun-google/> [consulta: 15 junio 2022].

ZAMBRANO, Lilia M., et al., 2021. Sistemas para la gerencia de ventas aplicados a la gestión en restaurantes. *Revista Ibérica De Sistemas e Tecnologías De Informação* [en línea]. s. 101-111. ISSN 16469895. Disponible en: <https://www.proquest.com/docview/2483102534/7A481E5CD8F34D73PQ/5?accountid=37408>

ANEXOS

ANEXO N° 01: Operacionalización de variables

Variables	Definición conceptual	Definición operacional	Dimensión	Indicador	Instrumento	Fórmula	Escala de medición
Sistema de Análisis Predictivo	El análisis predictivo es una ventaja principal que ofrece Machine Learning ya que utiliza los métodos de análisis de datos para encontrar comportamientos o relaciones de un conjunto de información. Guerra et al.(2020)	Permite reconocer un problema y dar solución a través de la tecnología	datos de entrada				
			algoritmo clasificación				
			resultados				
Gestión de ventas	“La Gestión de Ventas tiene como objetivo principal mantener y aumentar las ventas mediante la gestión del talento humano en otras palabras traducir los objetivos y estrategias de la empresa en objetivos y estrategias de recursos, lo cual se logra mediante la planeación estratégica de ventas además del mercado potencial de clientes”. Zambrano, et al.(2021)	Permite reconocer y predecir los productos que más se venden.	Rotación de productos	índice de rotación de productos	ficha de registro	$RP = \text{Ventas acumuladas} / \text{promedio del inventario}$ ALVAREZ Y PARADA(2020, p57)	razón
			Rentabilidad de productos	índice de ganancia de los productos	ficha de registro	$U = (PV_u - CV_u)$ PACHECO (2019, p28)	razón
			Precisión de ventas	índice de precisión de productos	Ficha de registro	$PVT = (1 - (P - V / V)) * 100$ P = Valor predicho V= Valor vendido	razón

Fuente: Elaboración propia

ANEXO N° 02: Matriz de consistencia

Problema	Objetivos	Hipótesis	Variables				Metodología
General	General	General	Independiente	Dimensiones	Indicadores	Técnica/ Instrumento	Tipo de investigación: Aplicada. Diseño de investigación: Experimental Población: 212 registros de venta Método de Investigación: Recolección de datos: Técnica: Fichaje Instrumento: Ficha de registro
¿Cómo influye el sistema de análisis predictivo para la gestión de ventas de la empresa TAURO A&F SAC?	Determinar la influencia del sistema de análisis predictivo para la gestión de ventas de la empresa TAURO A&F SAC	El sistema de análisis predictivo mejora la gestión de ventas en la empresa TAURO A&F SAC	Sistema de Análisis Predictivo				
Específicos	Específicos	Específicos	Dependiente				
1-¿Cómo influye el sistema de análisis predictivo en incrementar el índice de rotación de productos para la gestión de ventas de la empresa TAURO A&F SAC?	1- Determinar la influencia del sistema de análisis predictivo para incrementar el índice de rotación de productos en la gestión de ventas de la empresa TAURO A&F SAC	1) El sistema de análisis predictivo incremento el índice de rotación de productos para la gestión de ventas de la empresa TAURO A&F SAC	Gestión de ventas	Rotación de productos	índice de rotación de productos	Ficha de registro	
2- ¿Cómo influye el sistema de análisis predictivo en el índice de ganancia de los productos para la gestión de ventas de la empresa TAURO A&F SAC?	2- Determinar la influencia del sistema de análisis predictivo para incrementar el índice de ganancia de los productos en la gestión de ventas de la empresa TAURO A&F SAC	2) El sistema de análisis predictivo incremento el índice de ganancia de los productos para la gestión de ventas de la empresa TAURO A&F SAC		Rentabilidad de productos	índice de ganancia de los productos	Ficha de registro	
2- ¿Cómo influye el sistema de análisis predictivo en el índice de precisión de productos para la gestión de ventas de la empresa TAURO A&F SAC?	2- Determinar la influencia del sistema de análisis predictivo para incrementar el índice de precisión de productos en la gestión de ventas de la empresa TAURO A&F SAC	2) El sistema de análisis predictivo incremento el índice de precisión de productos para la gestión de ventas de la empresa TAURO A&F SAC		Precisión de ventas	índice de precisión de productos	Ficha de registro	

Fuente: Elaboración propia

ANEXO N° 03: Tabla de comparación entre metodologías de desarrollo

Metodologías para el desarrollo de proyectos de minería de datos	Entidad creadora	Autores	Fecha de creación	Tipos de proyectos a los que se ha aplicado	Fases / Etapas
KDD (Knowledge Discovery in Databases)	Fayyad	Hernando Camargo y Mario Silva	1996	Aplicación de una metodología adaptada de minería de datos en información del sector público.	Fases: 1. Selección 2. Pre-procesamiento / Limpieza 3. Transformación / Reducción 4. Minería de datos 5. Interpretación / Evaluación
SEMMA (Sample, Explore, Modify, Model y Assess)	SAS Institute	Efraín Oviedo, Ana Oviedo y Gloria Vélez	2012	Metodología de estudio de rendimiento académico mediante la minería de datos.	Fases: 1.Muestreo 2.Exploración 3.Manipulación 4.Modelado 5.Valoración
CRISP-DM (Cross Industry Standard Process for Data Mining)	SPSS, NCR y Daimler Chrysler	Juan Giraldo y Jovani Jiménez	1999	Metodología aplicada en una organización hospitalaria especializada en oncología que pretende mejorar la precisión en el diagnóstico de células cancerígenas.	Fases: 1. Comprensión del negocio 2. Comprensión de los datos 3. Preparación de los datos 4. Modelado 5. Evaluación 6. Despliegue

Fuente: Elaboración propia

ANEXO N° 04: Consentimiento informado

Anexo 1: Consentimiento informado

CONSENTIMIENTO INFORMADO

Yo Katherine Ovalle Pillco identificado(a) con DNI (carné de extranjería o pasaporte para extranjeros) N.º 74890821 he sido informado(a) sobre el procedimiento de la investigación titulada " SISTEMA DE ANÁLISIS PREDICTIVO PARA LA GESTIÓN DE VENTAS DE LA EMPRESA TAURO A&F SAC", cuyos autores son **ARAGON FARFAN CHRISTIAN JESUS JUNIOR** con DNI **43318657** y **SANCHEZ ALTAMIRANO CLEIVER DERMALI** con DNI **71396673** y se me ha entregado una copia de este consentimiento informado, fechado y firmado.

Además, se me ha explicado las características y el objetivo del estudio, así como los posibles beneficios de este. He contado con el tiempo y la oportunidad para realizar preguntas y plantear las dudas que poseía. Todas las preguntas fueron respondidas a mi entera satisfacción.

Se me ha asegurado que se mantendrá la confidencialidad de mis datos. Mi consentimiento lo otorgo de manera voluntaria y sé que soy libre de retirarme del estudio en cualquier momento, por cualquier razón de fuerza mayor. Por lo tanto, en forma consciente y voluntaria doy mi consentimiento para ser parte de esta investigación.

Lima, 17/10/2022

Katherine Ovalle Pillco

Apellidos y nombres

Firma

Huella



74890821

DNI

23

Edad

FEMENINO

Sexo (F:Femenino /M:Masculino)

ANEXO N° 05: Autorización para divulgación de resultados

Anexo 1: Autorización para la realización y difusión de resultados de la investigación

AUTORIZACIÓN PARA LA REALIZACIÓN Y DIFUSIÓN DE RESULTADOS DE LA INVESTIGACIÓN

Por medio del presente documento, Yo KATHERINE OVALLE PILLCO identificado con DNI N° 74890821 y representante legal de TAURO A&F SAC autorizo a CHRISTIAN JESÚS JUNIOR ARAGÓN FARFAN identificado con DNI N° 433186557 Y CLEIVER DERMALI SANCHEZ ALTAMIRANO con DNI N° 71396673 a realizar la investigación titulada: "SISTEMA DE ANÁLISIS PREDICTIVO PARA LA GESTIÓN DE VENTAS DE LA EMPRESA TAURO A&F SAC" y a difundir los resultados de la investigación utilizando el nombre de TAURO A&F SAC

Lima, 17 de Octubre de 2022

FIRMA



KATHERINE OVALLE PILLCO

DNI N° 74890821

ADMINISTRADOR

TAURO A&F SAC

ANEXO N° 06: Fichas de registro de las variables – Rotación de productos
Indicador: Rotación de productos.

Ficha de Registro					
Investigador	Aragón Farfan, Christian Jesús Junior Sanchez Altamirano, Cleiver Dermalí			TIPO DE PRUEBA	PRE TEST
Empresa	Tauro A&F SAC				
Variable	Gestión de ventas				
Dimensión	Rotación de productos				
Periodo	01/01/2022 – 30/01/2022				
INDICADOR	DESCRIPCIÓN	TÉCNICA	UNIDAD DE MEDIDA	FÓRMULA	
Rotación de productos	El objetivo principal es calcular el índice de rotación de productos	FICHAJE	Unidad	$RP=VA/PI$ VA = Ventas acumuladas PI=Promedio del inventario	
ÍTEM	FECHA	Ventas acumuladas (VA)	Promedio del inventario (PI)	Índice de rotación de productos (RP)	
1	01/01/2022	0	5830	0.00	
2	02/01/2022	0	5830	0.00	
3	03/01/2022	0	5830	0.00	
4	04/01/2022	0	5830	0.00	
5	05/01/2022	0	5830	0.00	
6	06/01/2022	0	5830	0.00	
7	07/01/2022	0	5830	0.00	
8	08/01/2022	550	5830	0.09	
9	09/01/2022	0	5830	0.00	
10	10/01/2022	330	5830	0.06	
11	11/01/2022	110	5830	0.02	
12	12/01/2022	440	5830	0.08	
13	13/01/2022	330	5830	0.06	
14	14/01/2022	0	5830	0.00	
15	15/01/2022	0	5830	0.00	
16	16/01/2022	0	5830	0.00	
17	17/01/2022	0	5830	0.00	
18	18/01/2022	0	5830	0.00	
19	19/01/2022	0	5830	0.00	
20	20/01/2022	0	5830	0.00	
21	21/01/2022	0	5830	0.00	
22	22/01/2022	0	5830	0.00	
23	23/01/2022	0	5830	0.00	
24	24/01/2022	0	5830	0.00	
25	25/01/2022	0	5830	0.00	
26	26/01/2022	0	5830	0.00	
27	27/01/2022	0	5830	0.00	
28	28/01/2022	0	5830	0.00	
29	29/01/2022	220	5830	0.04	
30	30/01/2022	0	5830	0.00	

Ficha de Registro					
Investigador	Aragón Farfan, Christian Jesús Junior Sanchez Altamirano, Cleiver Dermalí			TIPO DE PRUEBA	POST TEST
Empresa	Tauro A&F SAC				
Variable	Gestión de ventas				
Dimensión	Rotación de productos				
Periodo	01/06/2022 – 30/06/2022				
INDICADOR	DESCRIPCIÓN	TÉCNICA	UNIDAD DE MEDIDA	FÓRMULA	
Rotación de productos	El objetivo principal es calcular el índice de rotación de productos	FICHAJE	Unidad	$RP=VA/PI$ VA = Ventas acumuladas PI=Promedio del inventario	
ÍTEM	FECHA	Ventas acumuladas (VA)	Promedio del inventario (PI)	Índice de rotación de productos (RP)	
1	01/06/2022	110	5830	0.02	
2	02/06/2022	0	5830	0.00	
3	03/06/2022	0	5830	0.00	
4	04/06/2022	220	5830	0.04	
5	05/06/2022	220	5830	0.04	
6	06/06/2022	220	5830	0.04	
7	07/06/2022	330	5830	0.06	
8	08/06/2022	110	5830	0.02	
9	09/06/2022	110	5830	0.02	
10	10/06/2022	110	5830	0.02	
11	11/06/2022	220	5830	0.04	
12	12/06/2022	220	5830	0.04	
13	13/06/2022	220	5830	0.04	
14	14/06/2022	330	5830	0.06	
15	15/06/2022	110	5830	0.02	
16	16/06/2022	110	5830	0.02	
17	17/06/2022	110	5830	0.02	
18	18/06/2022	110	5830	0.02	
19	19/06/2022	220	5830	0.04	
20	20/06/2022	220	5830	0.04	
21	21/06/2022	330	5830	0.06	
22	22/06/2022	110	5830	0.02	
23	23/06/2022	110	5830	0.02	
24	24/06/2022	110	5830	0.02	
25	25/06/2022	110	5830	0.02	
26	26/06/2022	220	5830	0.04	
27	27/06/2022	110	5830	0.02	
28	28/06/2022	330	5830	0.06	
29	29/06/2022	110	5830	0.02	
30	30/06/2022	110	5830	0.02	

ANEXO N° 07: Fichas de registro de las variables – Ganancia de productos
Indicador: Ganancia de productos.

Ficha de Registro						
Investigador		Aragón Farfan, Christian Jesús Junior Sanchez Altamirano, Cleiver Dermal			TIPO DE PRUEBA	PRE TEST
Empresa		Tauro A&F SAC				
Variable		Gestión de ventas				
Dimensión		Rentabilidad de productos				
Periodo		01/01/2022– 30/01/2022				
INDICADOR		DESCRIPCIÓN		TÉCNICA	UNIDAD DE MEDIDA	FORMULA
Ganancia de los productos		El objetivo principal es calcular el índice de ganancia de productos		FICHAJE	Unidad	$U = PV_u - CV_u$ PV _u = Precio de venta CV _u = Precio de costo
N°	FECHA	Cantidad	Precio de costo (CVu)	Precio de venta (PVu)	Ganancia (U)	
1	01/01/2022	0	0	0	0	
2	02/01/2022	0	0	0	0	
3	03/01/2022	0	0	0	0	
4	04/01/2022	0	0	0	0	
5	05/01/2022	0	0	0	0	
6	06/01/2022	0	0	0	0	
7	07/01/2022	0	0	0	0	
8	08/01/2022	5	467.5	550	82.5	
9	09/01/2022	0	0	0	0	
10	10/01/2022	3	280.5	330	49.5	
11	11/01/2022	1	93.5	110	16.5	
12	12/01/2022	4	374	440	66	
13	13/01/2022	3	280.5	330	49.5	
14	14/01/2022	0	0	0	0	
15	15/01/2022	0	0	0	0	
16	16/01/2022	0	0	0	0	
17	17/01/2022	0	0	0	0	
18	18/01/2022	0	0	0	0	
19	19/01/2022	0	0	0	0	
20	20/01/2022	0	0	0	0	
21	21/01/2022	0	0	0	0	
22	22/01/2022	0	0	0	0	
23	23/01/2022	0	0	0	0	
24	24/01/2022	0	0	0	0	
25	25/01/2022	0	0	0	0	
26	26/01/2022	0	0	0	0	
27	27/01/2022	0	0	0	0	
28	28/01/2022	0	0	0	0	
29	29/01/2022	2	187	220	33	
30	30/01/2022	0	0	0	0	

Ficha de Registro

Investigador		Aragón Farfan, Christian Jesús Junior Sanchez Altamirano, Cleiver Dermalí			TIPO DE PRUEBA	POST TEST
Empresa		Tauro A&F SAC				
Variable		Gestión de ventas				
Dimensión		Rentabilidad de productos				
Periodo		01/06/2022 – 30/06/2022				
INDICADOR		DESCRIPCIÓN		TÉCNICA	UNIDAD DE MEDIDA	FORMULA
Ganancia de los productos		El objetivo principal es calcular el índice de ganancia de productos		FICHAJE	Unidad	$U = PV_u - CV_u$ PV _u = Precio de venta CV _u = Precio de costo
N°	FECHA	Cantidad	Precio de costo (CV _u)	Precio de venta (PV _u)	Ganancia (U)	
1	01/06/2022	1	93.5	110	16.5	
2	02/06/2022	0	0	0	0	
3	03/06/2022	0	0	0	0	
4	04/06/2022	2	187	220	33	
5	05/06/2022	2	187	220	33	
6	06/06/2022	2	187	220	33	
7	07/06/2022	3	280.5	330	49.5	
8	08/06/2022	1	93.5	110	16.5	
9	09/06/2022	1	93.5	110	16.5	
10	10/06/2022	1	93.5	110	16.5	
11	11/06/2022	2	187	220	33	
12	12/06/2022	2	187	220	33	
13	13/06/2022	2	187	220	33	
14	14/06/2022	3	280.5	330	49.5	
15	15/06/2022	1	93.5	110	16.5	
16	16/06/2022	1	93.5	110	16.5	
17	17/06/2022	1	93.5	110	16.5	
18	18/06/2022	1	93.5	110	16.5	
19	19/06/2022	2	187	220	33	
20	20/06/2022	2	187	220	33	
21	21/06/2022	3	280.5	330	49.5	
22	22/06/2022	1	93.5	110	16.5	
23	23/06/2022	1	93.5	110	16.5	
24	24/06/2022	1	93.5	110	16.5	
25	25/06/2022	1	93.5	110	16.5	
26	26/06/2022	2	187	220	33	
27	27/06/2022	1	93.5	110	16.5	
28	28/06/2022	3	280.5	330	49.5	
29	29/06/2022	1	93.5	110	16.5	
30	30/06/2022	1	93.5	110	16.5	

ANEXO N° 08: Fichas de registro de las variables – Precisión de productos
Indicador: Precisión de productos.

Ficha de Registro					
Investigador		Aragón Farfan, Christian Jesús Junior Sanchez Altamirano, Cleiver Dermalí		TIPO DE PRUEBA	PRE TEST
Empresa		Tauro A&F SAC			
Variable		Gestión de ventas			
Dimensión		Precisión de productos			
Periodo		01/01/2022– 31/01/2022			
INDICADOR		DESCRIPCIÓN	TÉCNICA	UNIDAD DE MEDIDA	FÓRMULA
Precisión de productos (PVT)		El objetivo principal es calcular el índice de precisión de productos en lo referido a las ventas reales y predichas	FICHAJE	Unidad	$PVT = \frac{1 - (P - V /V)}{1} * 100$ P = Valor predicho V = Valor vendido
ÍTEM	FECHA	Valor predicho(P)	Valor vendido(V)		Precisión de venta(PVT)
1	01/01/2022	0	0		0.00
2	02/01/2022	0	0		0.00
3	03/01/2022	0	0		0.00
4	04/01/2022	1	0		0.00
5	05/01/2022	1	0		0.00
6	06/01/2022	1	0		0.00
7	07/01/2022	2	0		0.00
8	08/01/2022	4	5		0.80
9	09/01/2022	0	0		0.00
10	10/01/2022	5	3		0.33
11	11/01/2022	3	1		-1.00
12	12/01/2022	3	4		0.75
13	13/01/2022	4	3		0.67
14	14/01/2022	1	0		0.00
15	15/01/2022	1	0		0.00
16	16/01/2022	2	0		0.00
17	17/01/2022	0	0		0.00
18	18/01/2022	1	0		0.00
19	19/01/2022	1	0		0.00
20	20/01/2022	0	0		0.00
21	21/01/2022	2	0		0.00
22	22/01/2022	3	0		0.00
23	23/01/2022	0	0		0.00
24	24/01/2022	0	0		0.00
25	25/01/2022	1	0		0.00
26	26/01/2022	1	0		0.00
27	27/01/2022	0	0		0.00
28	28/01/2022	0	0		0.00
29	29/01/2022	4	2		0.00
30	30/01/2022	0	0		0.00

Ficha de Registro					
Investigador	Aragón Farfan, Christian Jesús Junior Sanchez Altamirano, Cleiver Dermalí			TIPO DE PRUEBA	POST TEST
Empresa	Tauro A&F SAC				
Variable	Gestión de ventas				
Dimensión	Precisión de productos				
Periodo	01/06/2022– 31/06/2022				
INDICADOR	DESCRIPCIÓN	TÉCNICA	UNIDAD DE MEDIDA	FÓRMULA	
Precisión de productos (PVT)	El objetivo principal es calcular el índice de precisión de productos en lo referido a las ventas reales y predichas	FICHAJE	Unidad	$PVT = \frac{1 - (P - V /V)}{100}$ P = Valor predicho V = Valor vendido	
ÍTEM	FECHA	Valor predicho(P)	Valor vendido(V)	Precisión de venta(PVT)	
1	01/06/2022	1	0	0.00	
2	02/06/2022	0	0	0.00	
3	03/06/2022	0	4	0.00	
4	04/06/2022	2	0	0.00	
5	05/06/2022	2	0	0.00	
6	06/06/2022	2	4	0.50	
7	07/06/2022	3	7	0.43	
8	08/06/2022	1	13	0.08	
9	09/06/2022	1	0	0.00	
10	10/06/2022	1	3	0.33	
11	11/06/2022	2	0	0.00	
12	12/06/2022	2	0	0.00	
13	13/06/2022	2	0	0.00	
14	14/06/2022	3	0	0.00	
15	15/06/2022	1	1	1.00	
16	16/06/2022	1	4	0.25	
17	17/06/2022	1	0	0.00	
18	18/06/2022	1	2	0.50	
19	19/06/2022	2	0	0.00	
20	20/06/2022	2	0	0.00	
21	21/06/2022	3	5	0.60	
22	22/06/2022	1	9	0.11	
23	23/06/2022	1	0	0.00	
24	24/06/2022	1	0	0.00	
25	25/06/2022	1	4	0.25	
26	26/06/2022	2	0	0.00	
27	27/06/2022	1	0	0.00	
28	28/06/2022	3	0	0.00	
29	29/06/2022	1	0	0.00	
30	30/06/2022	1	0	0.00	

ANEXO N° 09: Certificado de validez de contenido del instrumento que mide

CERTIFICADO DE VALIDEZ DE CONTENIDO DEL INSTRUMENTO QUE MIDE:

N°	DIMENSIONES / ítems	Pertinencia ¹		Relevancia ²		Claridad ³		Sugerencias
		Si	No	Si	No	Si	No	
1	INDICADOR: índice de rotación de productos $RP = \frac{\text{Ventas acumuladas}}{\text{Promedio de inventario}}$	X		X		X		
2	INDICADOR: índice de ganancia de los productos $U = PVu - CVu$	X		X		X		
3	INDICADOR: índice de precisión de productos $PVT = \left(1 - \frac{ P-V }{V}\right) * 100$	X		X		X		

Observaciones (precisar si hay suficiencia): _____

Opinión de aplicabilidad: **Aplicable [x]** **Aplicable después de corregir []** **No aplicable []**

Apellidos y nombres del juez validador: Mendoza Apaza Fernando

DNI: 10363032

Especialidad del validador: Magister en Administración, Ingeniero de Sistemas

¹Pertinencia: El ítem corresponde al concepto teórico formulado.

²Relevancia: El ítem es apropiado para representar al componente o dimensión específica del constructo

³Claridad: Se entiende sin dificultad alguna el enunciado del ítem, es conciso, exacto y directo

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

20 de noviembre del 2022



Firma del Experto Informante.

ANEXO N° 10: Validación de la metodología de desarrollo

VALIDACIÓN DE LA METODOLOGÍA DE DESARROLLO

Apellidos y Nombres del Experto:

Mendoza Apaza Fernando

Título y/o Grado Académico:

Ingeniero Electrónico / Dr. En Educación

Doctor (x) Magister () Ingeniero (x) Licenciado () Otro ()

Fecha: 20/11/2022

Título de Investigación: Sistema de análisis predictivo para la gestión de ventas de la empresa TAURO A&F

Autores:

Aragón Farfan, Christian Jesús Junior

MUY MAL (1) MALO (2) REGULAR (3) BUENO (4) EXCELENTE (5)

ÍTEM	PREGUNTAS	METODOLOGÍA		
		CRISP-DM	SEMMA	KDD
1	¿Qué metodología es la más adecuada para este tipo de investigación?	5		
2	¿Qué metodología define las necesidades del cliente?	4		
3	¿Qué metodología estudia y comprende los datos?	4		
4	¿Qué metodología realiza un análisis de datos y selecciona características?	4		
5	¿Qué metodología logra realizar un modelado con los datos obtenidos?	4		
6	¿Qué metodología evalúa los resultados obtenidos?	5		
7	¿Qué metodología realiza un despliegue con los datos obtenidos?	5		
PUNTUACIÓN		BUENO		

SUGERENCIAS

FIRMA DEL EXPERTO

ANEXO N° 11: Validación por juicio de expertos: rotación de productos

VALIDACIÓN DEL INSTRUMENTO DE EXPERTOS: índice de rotación de productos

I. DATOS GENERALES

Apellidos y Nombres del Experto: Mendoza Apaza Fernando
 Título y/o Grado Académico: Ingeniero Electrónico / Dr. En Educación

Doctor (x) Magister () Ingeniero (x) Licenciado () Otro ().....

Universidad que labora: Universidad César Vallejo
 Fecha: 20/11/2022

Título de Investigación: Sistema de análisis predictivo para la gestión de ventas de la empresa TAURO A&F

Autores:

- Aragón Farfan, Christian Jesús Junior

Deficiente (0-20%) Regular(21-50%) Bueno(51-70%) Muy Bueno(71-80%) Excelente(81-100%)

II. ASPECTOS DE VALIDACION

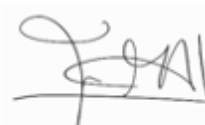
INDICADOR	CRITERIO	VALORACION				
		0-20%	21-50%	51-70%	71-80%	81-100%
CLARIDAD	Es formulado con lenguaje apropiado.				80%	
OBJETIVIDAD	Está expresado en conducta observable.				80%	
ACTUALIDAD	Es adecuado el avance, la ciencia y tecnología.				80%	
ORGANIZACION	Existe una organización lógica.				80%	
SUFICIENCIA	Comprende los aspectos de cantidad y calidad.				80%	
INTENCIONALIDAD	Adecuado para valorar los aspectos del sistema metodológico y científico.				80%	
CONSISTENCIA	Está basado en aspectos teóricos y científicos.				80%	
COHERENCIA	En los datos respecto al indicador.				80%	
METODOLOGIA	Responde al propósito de investigación.				80%	
PERTENENCIA	El instrumento es adecuado al tipo de investigación.				80%	
TOTAL					80%	

III. PROMEDIO DE VALIDACION

80%

IV. OPCION DE APLICABILIDAD

- (x) El instrumento puede ser aplicado, tal como está elaborado
 El instrumento debe ser mejorado antes de ser
 () aplicado



FIRMA DEL EXPERTO

ANEXO N° 12: Validación por juicio de expertos: ganancia de los productos

TABLA DE VALIDACIÓN DEL INSTRUMENTO DE EXPERTOS: índice de ganancia de los productos

I. DATOS GENERALES

Apellidos y Nombres del Experto: Mendoza Apaza Fernando
 Título y/o Grado Académico: Ingeniero Electrónico / Dr. En Educación

Doctor (x) Magister () Ingeniero (x) Licenciado () Otro ().....

Universidad que labora: Universidad César Vallejo
 Fecha: 20/11/2022

Título de Investigación: Sistema de análisis predictivo para la gestión de ventas de la empresa TAURO A&F

Autores:

- Aragón Farfan, Christian Jesús Junior

Deficiente (0-20%) Regular (21-50%) Bueno (51-70%) Muy Bueno (71-80%) Excelente (81-100%)

II. ASPECTOS DE VALIDACION

INDICADOR	CRITERIO	VALORACION				
		0-20%	21-50%	51-70%	71-80%	81-100%
CLARIDAD	Es formulado con lenguaje apropiado.				80%	
OBJETIVIDAD	Esta expresado en conducta observable.				80%	
ACTUALIDAD	Es adecuado el avance, la ciencia y tecnología.				80%	
ORGANIZACION	Existe una organización lógica.				80%	
SUFICIENCIA	Comprende los aspectos de cantidad y calidad.				80%	
INTENCIONALIDAD	Adecuado para valorar los aspectos del sistema metodológico y científico.				80%	
CONSISTENCIA	Está basado en aspectos teóricos y científicos.				80%	
COHERENCIA	En los datos respecto al indicador.				80%	
METODOLOGIA	Responde al propósito de investigación.				80%	
PERTENENCIA	El instrumento es adecuado al tipo de investigación.				80%	
TOTAL					80%	

III. PROMEDIO DE VALIDACIONx

80%

IV. OPCION DE APLICABILIDAD

- (x) El instrumento puede ser aplicado, tal como está elaborado
 El instrumento debe ser mejorado antes de ser
 () aplicado



FIRMA DEL EXPERTO

ANEXO N° 13: Validación por juicio de expertos: precisión de productos

TABLA DE VALIDACIÓN DEL INSTRUMENTO DE EXPERTOS: índice de precisión de productos

I. DATOS GENERALES

Apellidos y Nombres del Experto: Mendoza Apaza Fernando
 Título y/o Grado Académico: Ingeniero Electrónico / Dr. En Educación

Doctor (x) Magister () Ingeniero (x) Licenciado () Otro ().....

Universidad que labora: Universidad César Vallejo
 Fecha: 20/11/2022

Título de Investigación: Sistema de análisis predictivo para la gestión de ventas de la empresa TAURO A&F

Autores:

- Aragón Farfan, Christian Jesús Junior

Deficiente (0-20%) Regular (21-50%) Bueno(51-70%) Muy Bueno (71-80%) Excelente (81-100%)

II. ASPECTOS DE VALIDACION

INDICADOR	CRITERIO	VALORACION				
		0-20%	21-50%	51-70%	71-80%	81-100%
CLARIDAD	Es formulado con lenguaje apropiado.				80%	
OBJETIVIDAD	Esta expresado en conducta observable.				80%	
ACTUALIDAD	Es adecuado el avance, la ciencia y tecnología.				80%	
ORGANIZACION	Existe una organización lógica.				80%	
SUFICIENCIA	Comprende los aspectos de cantidad y calidad.				80%	
INTENCIONALIDAD	Adecuado para valorar los aspectos del sistema metodológico y científico.				80%	
CONSISTENCIA	Está basado en aspectos teóricos y científicos.				80%	
COHERENCIA	En los datos respecto al indicador.				80%	
METODOLOGIA	Responde al propósito de investigación.				80%	
PERTENENCIA	El instrumento es adecuado al tipo de investigación.				80%	
TOTAL					80%	

III. PROMEDIO DE VALIDACION

80%

IV. OPCION DE APLICABILIDAD

- (x) El instrumento puede ser aplicado, tal como está elaborado
 El instrumento debe ser mejorado antes de ser
 () aplicado



FIRMA DEL EXPERTO

ANEXO N° 14: Solicitud de cambio de título

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS RESOLUCIÓN DIRECTORAL N° 0346- 2022-EP-ING-SIS.UCV LIMA NORTE

Los Olivos, 18 de Noviembre del 2022

VISTO: El Dictamen N° 0240-2022-EP-ING-SIS.UCV LIMA NORTE de fecha 11 de Noviembre, presentado por la comisión evaluadora de la Tesis designado por **coordinación de escuela** de Ingeniería de Sistemas, en el cual se establece la procedencia para el cambio de título de la Tesis "**SISTEMA DE ANÁLISIS PREDICTIVO PARA LA GESTIÓN DE VENTAS DE LA EMPRESA TAURO A&F SAC**" por el (los) estudiante (s) **CHRISTIAN JESUS JUNIOR ARAGON FARFAN**.

CONSIDERANDO

Que, mediante **RESOLUCIÓN DIRECTORAL N° 0293-2022-EP-ING-SIS.UCV LIMA NORTE** de fecha 04 de Noviembre del 2022, se designó a la comisión Evaluadora de la Tesis "**SISTEMA DE ANÁLISIS PREDICTIVO PARA LA GESTIÓN DE VENTAS DE LA EMPRESA TAURO A&F SAC**" (la)estudiante **CHRISTIAN JESUS JUNIOR ARAGON FARFAN**; a los siguientes docentes:

Dr. YOHAN ROY ALARCON CAJAS
Dr. JORGE ISAAC NECOCHEA CHAMORRO
Dra. YESENIA DEL ROSARIO VÁSQUEZ VALENCIA

Estando a lo expuesto y en uso de las atribuciones conferidas y de conformidad con las normas y reglamentos vigentes;

SE RESUELVE

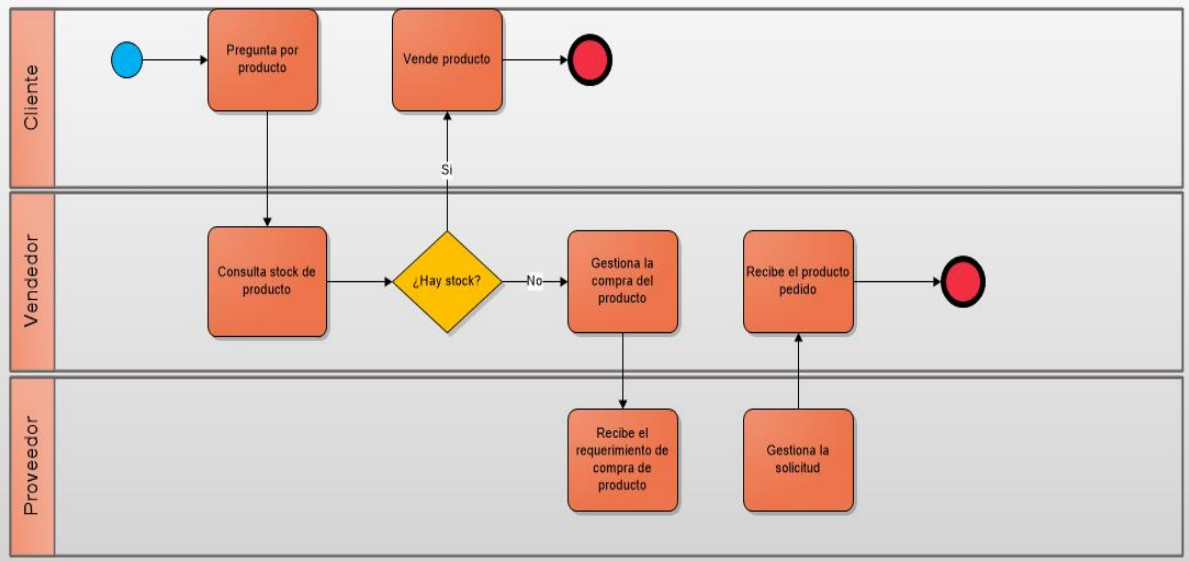
ARTÍCULO 1º: SE APROBO EL CAMBIO DE TITULO de la Tesis denominada: "**SISTEMA DE ANÁLISIS PREDICTIVO PARA LA GESTIÓN DE VENTAS DE LA EMPRESA TAURO A&F SAC**" Presentada por el (los) estudiante (s) **CHRISTIAN JESUS JUNIOR ARAGON FARFAN**.

Regístrese, comuníquese y archívese.



Dra. YESENIA DEL ROSARIO VÁSQUEZ VALENCIA
Coordinadora Académica
Escuela Profesional de Ingeniería de Sistemas
UCV Lima Norte

ANEXO N° 15: Flujoograma de gestión de venta



ANEXO N° 16: Tecnologías utilizadas

Las tecnologías usadas son las siguientes:

- Python 3.2 : Lenguaje usado en el desarrollo del modelo
- Visual Studio Code: Editor de código
- Google Colaboratory: Utilitario para implementar modelos de machine learning

Procedemos a descargar e instalar la última versión de Python en Visual Studio Code

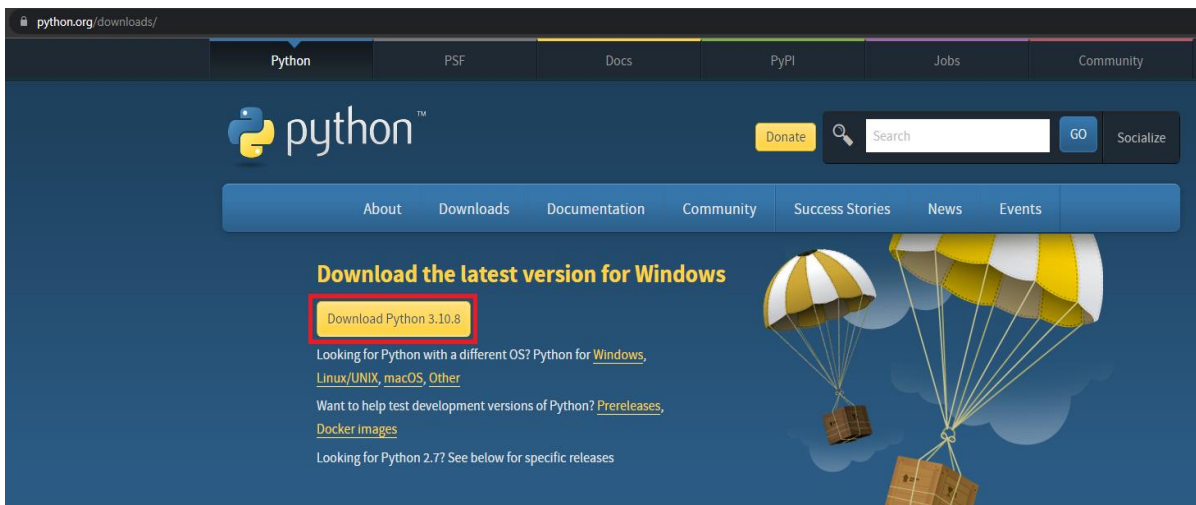


Figura 7: Lenguaje de programación para el desarrollo del modelo (Python 2022)

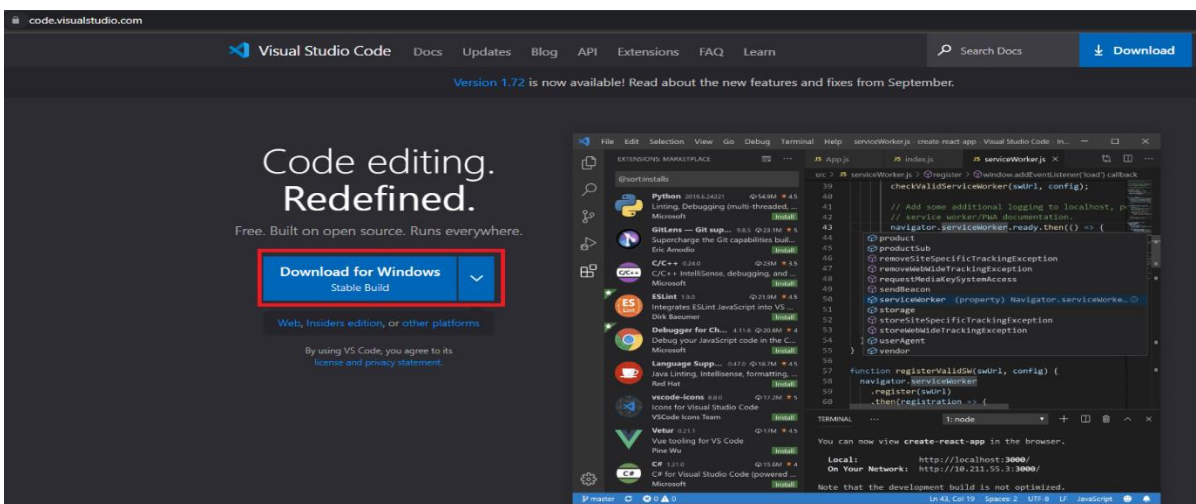
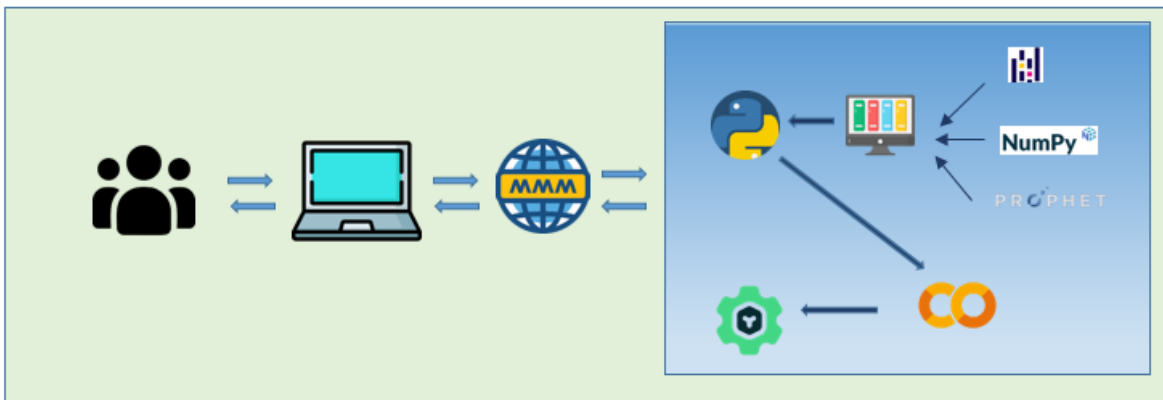


Figura 8: Visual Studio Code, editor de código (Visual Studio Code)



Figura 9: Google Colaboratory, utilitario para implementación de modelos de Machine learning (Google Colaboratory)

ANEXO N° 17: Arquitectura tecnológica para el desarrollo del modelo



ANEXO N° 18: Arquitectura tecnológica para el usuario final

En la siguiente imagen se muestra como interactúa el usuario con la computadora para obtener datos del modelo desarrollado.



ANEXO N° 19: Aplicación de la metodología CRISP-DM al modelo

1. Comprensión del negocio

En esta etapa se realiza el entendimiento de los objetivos y requerimientos del negocio para plantearlos en el proyecto.

Determinar los objetivos del negocio

El objetivo de aplicar Machine Learning en la organización es poder predecir la cantidad de productos a vender.

- Predecir la cantidad de ventas mensuales por producto.
- Tener el stock adecuado de acuerdo a lo analizado.

Evaluación de la situación

Se tiene el histórico de ventas de productos desde noviembre 2021 a Mayo 2022.en un archivo Excel.

Inventario de recursos

La implementación del modelo se hará con el lenguaje Python, no es necesario una base de datos. Los recursos de hardware tienen las siguientes características:

Tabla 11: Recursos tecnológicos para el desarrollo del modelo

Hardware	Características	Cantidad
Computadora portátil	HP- 15-dy2056la, Intel Core i5-1135G7, 8GB RAM-2666 MHz, Velocidad de procesador-3.6 GHz,	2

Fuente: Elaboración propia

Riesgos y contingencias

La integridad de los datos tiene que ser veraz para que los resultados del modelo tengan la precisión adecuada.

Determinar los objetivos del proyecto

Objetivos

Los objetivos son:

- Tener fiabilidad en la predicción de la cantidad de ventas mensuales por producto.
- Predecir las ganancias por los productos predichos.
- Predecir el índice de rotación por los productos predichos.

2. Comprensión de los datos

Recolección de los datos

La organización género reportes de ventas por mes en Excel los cuales se agruparon en un solo archivo, los campos que integran los reportes son:

Tabla 12: Campos del reporte de ventas

Campo
idarti
cdarti
dslart
caarti
Ganancia_total
Ganancia_unit
Prc_unit
Prc_vent
IGV
Prc_final
tpdocu
dstpdocu
nuseri
nudocu(B/F)
fedocu
idccte
dsrsoc

Fuente: Empresa Tauro A&F sac

Descripción de los datos

Tabla 13: Descripción de los datos

Campo	Descripción
idarti	Número del artículo en la base de datos
cdarti	Código de identificación del articulo
dslart	Descripción del articulo
caarti	Cantidad de ventas del articulo
Ganancia_total	Ganancias por cantidad de ventas
Ganancia_unit	Ganancia por unidad
Prc_unit	Precio por unidad
Prc_vent	Precio de venta
IGV	Impuesto
Prc_final	Precio de venta final
tpdocu	Tipo de documento de venta (boleta / factura)
dstpdocu	Descripción del tipo de documento de venta
nuseri	Serie interna del tipo de documento de venta
nudocu(B/F)	Serie del tipo de documento de venta
fedocu	Fecha de venta
idccte	Código interno
dsrsoc	Descripción del cliente

Fuente: Empresa Tauro A&F sac

Exploración de los datos

Para poder realizar un análisis previo de los datos se tuvo que consolidar los distintos reportes mensuales en un solo archivo Excel.

- Se tienen 410 registros de venta del producto a analizar.

Verificar la calidad de los datos

Se verifica que hay vacíos en las fechas de ventas del producto, esto se produce porque no hay ventas diarias o fines de semana.

Id_articulo	Descripcion	Cantidad	Fch_venta
2853	BATERIA CAPSA 36I 650	1	10/11/2021
2853	BATERIA CAPSA 36I 650	1	10/11/2021
2853	BATERIA CAPSA 36I 650	1	11/11/2021
2853	BATERIA CAPSA 36I 650	2	13/11/2021
2853	BATERIA CAPSA 36I 650	10	13/11/2021
2853	BATERIA CAPSA 36I 650	5	13/11/2021
2853	BATERIA CAPSA 36I 650	3	15/11/2021
2853	BATERIA CAPSA 36I 650	10	15/11/2021
2853	BATERIA CAPSA 36I 650	1	16/11/2021
2853	BATERIA CAPSA 36I 650	2	23/11/2021
2853	BATERIA CAPSA 36I 650	2	27/11/2021

Figura 10: Verificación de la calidad de los datos (Elaboración propia)

3. Preparación de los datos

Seleccionar los datos

Para poder realizar un análisis de acuerdo a lo requerido se eligieron los siguientes campos:

Tabla 14: Selección de datos

Campo	Descripción	Nombre de campo modificado
idarti	Número del artículo en la base de datos	Producto
caarti	Cantidad de ventas del artículo	Ventas
fedocu	Fecha de venta	Fecha

Fuente: Elaboración propia

Además que se eligió el producto "BATERIA CAPSA 36I 650" como criterio de inclusión del modelo

Limpieza de datos

Para poder optimizar la calidad de los datos se procedió a agrupar las ventas por día.

fecha	producto	ventas
01/11/2021	2853	0
02/11/2021	2853	0
03/11/2021	2853	0
04/11/2021	2853	0
05/11/2021	2853	0
06/11/2021	2853	0
07/11/2021	2853	0
08/11/2021	2853	0
09/11/2021	2853	0
10/11/2021	2853	2
11/11/2021	2853	1
12/11/2021	2853	0

Figura 11: Limpieza de los datos (Elaboración propia)

Estructuración de los datos

Se agregó el campo “local” con valor de 1 que es el id de organización para poder realizar un mejor entendimiento.

fecha	local	producto	ventas
01/11/2021	1	2853	0
02/11/2021	1	2853	0
03/11/2021	1	2853	0
04/11/2021	1	2853	0
05/11/2021	1	2853	0
06/11/2021	1	2853	0
07/11/2021	1	2853	0
08/11/2021	1	2853	0
09/11/2021	1	2853	0
10/11/2021	1	2853	2
11/11/2021	1	2853	1

Figura 12: Estructuración de los datos (Elaboración propia)

Integración de los datos

Para poder optimizar la calidad de los datos se procedió a completar las fechas desde el primer y último registro de venta del mes.

Formateo de los datos

Para una mejor visibilidad y procesado de los datos se cambió el código del producto.

Id_articulo	id_art	Descripcion
2853	3	BATERIA CAPSA 36I 650

Figura 13: Formateo de los datos (Elaboración propia)

Quedando como resultado la siguiente imagen como input para el modelo

fecha	local	producto	ventas
01/11/2021	1	3	0
02/11/2021	1	3	0
03/11/2021	1	3	0
04/11/2021	1	3	0
05/11/2021	1	3	0
06/11/2021	1	3	0
07/11/2021	1	3	0
08/11/2021	1	3	0
09/11/2021	1	3	0
10/11/2021	1	3	2
11/11/2021	1	3	1
12/11/2021	1	3	0

Figura 14: Datos de entrada (Elaboración propia)

4. Modelamiento de los datos

Para iniciar con el desarrollo del modelo se utilizó Google colab que es un utilitario online de Google para poder desarrollar modelos de Machine Learning

Selección de técnica de modelado

Se realizó la prueba con las siguientes técnicas:

- **Mlp (multilayer perceptron)**
- **ElasticNet**
- **Prophet**

Por lo que en la imagen podemos notar que Prophet es la que obtiene mejores resultados, al ser un procedimiento para series temporales es lo ideal.

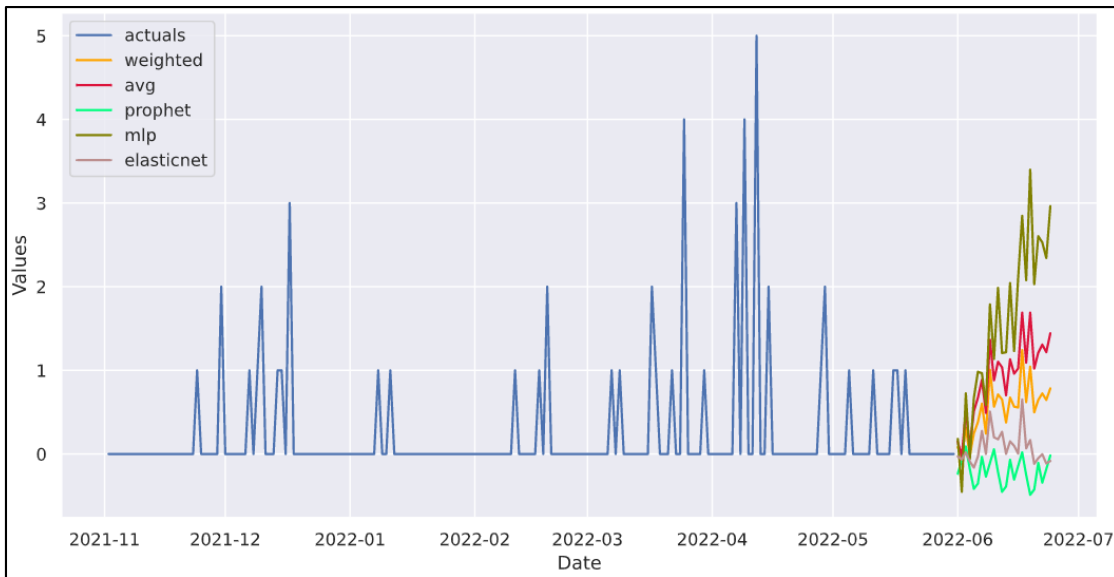


Figura 15: Técnicas de modelado (Elaboración propia)

Construcción del modelo

Librerías que se van a usar en el modelo:

```

1 #Libreria de procesamiento numerico
2 import numpy as np
3 #Libreria para analisis de datos
4 import pandas as pd
5 #Libreria para presentar los datos
6 import matplotlib.pyplot as plt
7 #Libreria para realizar la prediccion
8 from prophet import Prophet
9 import warnings

```

Figura 16: Librerías del modelo (Elaboración propia)

Se carga el archivo con los datos de entrada de los productos a un dataframe llamado df_train.

```

1 archivo = '/content/Dataset_ventas_modelado_proyecto.xlsx'
2 df_train = pd.read_excel(archivo, sheet_name='Análisis')

```

Figura 17: Lectura de los datos (Elaboración propia)

Se leen las columnas que integran el dataframe

```
1 # Lee las columnas
2 df_train.columns

Index(['fecha', 'local', 'producto', 'ventas'], dtype='object')
```

Figura 18: Lectura de las columnas (Elaboración propia)

Se cambia el nombre de las columnas para mejor entendimiento del código

```
1 # Cambia el nombre de las columnas solo por conveniencia de facilidad de código
2 df_train.rename (columns = {'fecha': 'date', 'local': 'store', 'producto': 'item', 'ventas': 'sales'}, inplace = True )
```

Figura 19: Renombramiento de las columnas (Elaboración propia)

Se selecciona como index la columna fecha, el cual es el valor que se le va asignar a cada fila para facilitar la búsqueda de información.

```
1 # Usar la fecha como indice de filas
2 df_train.index = pd.to_datetime(df_train['date'])
3 df_train.drop('date', axis=1, inplace=True)
```

Figura 20: Identificación del índice de las columnas (Elaboración propia)

Se muestran las primeras filas del dataframe

	store	item	sales
date			
2021-11-01	1	1	0
2021-11-02	1	1	0
2021-11-03	1	1	0
2021-11-04	1	1	0
2021-11-05	1	1	0
2021-11-06	1	1	0

Figura 21: Muestra de datos del dataframe (Elaboración propia)

Se verifica el rango de datos que contiene el dataset

```
1 # Solo para constatar el intervalo de datos que tenemos fecha minima y fecha maxima
2 df_fecha=pd.DataFrame() ;df_fecha['fecha']=df_train.index
3 print('Los datos empiezan el ', df_fecha.fecha.min() , 'y termina el ' , df_fecha.max())
4
```

Los datos empiezan el 2021-11-01 00:00:00 y termina el fecha 2022-05-31
dtype: datetime64[ns]

Figura 22: Rango de fechas del dataset (Elaboración propia)

Se validan los datos del producto a predecir

```
1 # Aqui se introduce el numero de producto a predecir
2 numero_producto=3
3 # Se valida que el numero este entre el intervalo correspondiente
4 if numero_producto in df_train['item'].value_counts():
5     print('El producto existe')
6     # vamos a filtrar los productos que estan nominados como numero 1 y obtenemos sus dimensiones
7     df_producto=df_train[df_train['item']==numero_producto]
8     print('La serie temporal de este producto es la siguiente')
9     print(df_producto.head(10))
10 else:
11     print(' El producto No existe')
12
```

El producto existe
La serie temporal de este producto es la siguiente

date	store	item	sales
2021-11-01	1	3	0
2021-11-02	1	3	0
2021-11-03	1	3	0
2021-11-04	1	3	0
2021-11-05	1	3	0
2021-11-06	1	3	0
2021-11-07	1	3	0
2021-11-08	1	3	0
2021-11-09	1	3	0
2021-11-10	1	3	2

Figura 23: Validación del producto a predecir (Elaboración propia)

Se muestra el comportamiento del producto a predecir.

```

1 # Graficar las ventas del producto para visualizar el comportamiento
2 plt.plot(df_producto.index,df_producto.sales , color='green')
3 plt.xlabel('fechas')
4 plt.ylabel('Cantidad vendida')
5 plt.title("Linear graph")
6 plt.show()

```

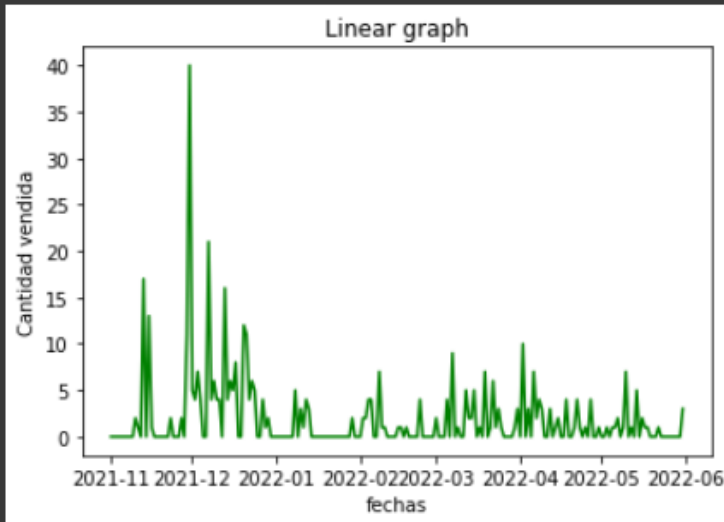


Figura 24: Comportamiento del producto a predecir (Elaboración propia)

Se le asignan a “x” los intervalos de tiempo y a “Y” las cantidades vendidas, además se preparan los datos de entrada al modelo teniendo un dataframe con las columnas ds (tiempo) e y para las ventas. Se tiene que tener en cuenta que para el modelo Prophet es siempre un marco de datos de dos columnas en este caso ds e y.

La columna ds (marca de fecha) debe tener un formato aceptado por la librería Pandas por lo cual se cambia a AAA-MM-DD o AAAA-MM-DD HH:MM:SS para una marca de tiempo.

La columna y debe ser numérica y va representar la medida a pronosticar.

```

1 # Asignamos a X los intervalos de tiempo y a Y las cantidades vendidas
2 X=df_producto.index;Y=df_producto.sales
3 # Preparamos los datos de entrada al modelo con un DataFrame que tenga las columnas ds(tiempo) e y para las ventas
4 # La entrada a Prophet es siempre un marco de datos con dos columnas: ds e y. La columna ds (marca de fecha) debe tener un formato esperado por Pandas,
5 # idealmente AAAA-MM-DD para una fecha o AAAA-MM-DD HH:MM:SS para una marca de tiempo.
6 # La columna y debe ser numérica y representa la medida que deseamos pronosticar.
7 df = pd.DataFrame()
8 df['ds']=X
9 Y1=[ y for y in df_producto.sales]
10 df['y']=Y1
11 df.head(10)

```

	ds	y
0	2021-11-01	0
1	2021-11-02	0
2	2021-11-03	0
3	2021-11-04	0
4	2021-11-05	0
5	2021-11-06	0
6	2021-11-07	0
7	2021-11-08	0
8	2021-11-09	0
9	2021-11-10	2

Figura 25: Columnas como entrada al modelo (Elaboración propia)

Se inicia el entrenamiento con Prophet

```

1 # Entrenamos 1 modelo prophet en Python
2 modelo = Prophet()
3 modelo.fit(df)

```

```

INFO:prophet:Disabling yearly seasonality. Run prophet with yearly_seasonality=True to override this.
INFO:prophet:Disabling daily seasonality. Run prophet with daily_seasonality=True to override this.
DEBUG:cmdstanpy:input tempfile: /tmp/tmpdz2_52wo/spazxm96.json
DEBUG:cmdstanpy:input tempfile: /tmp/tmpdz2_52wo/wd_eewob.json
DEBUG:cmdstanpy:idx 0
DEBUG:cmdstanpy:running CmdStan, num_threads: None
DEBUG:cmdstanpy:CmdStan args: ['/usr/local/lib/python3.7/dist-packages/prophet/stan_model/prophet_model.bin',
23:04:06 - cmdstanpy - INFO - Chain [1] start processing
INFO:cmdstanpy:Chain [1] start processing
23:04:06 - cmdstanpy - INFO - Chain [1] done processing
INFO:cmdstanpy:Chain [1] done processing
<prophet.forecaster.Prophet at 0x7f8c48086450>

```

Figura 26: Entrenamiento del modelo (Elaboración propia)

Se crea una función para poder usar el modelo en base a fechas.

```

1 # Aquí ingresas la fecha a predecir
2
3 def prediccion_ventas(fecha_evaluar,modelo):
4     d=pd.DataFrame()
5     d['ds']=[fecha_evaluar]
6     predictor = modelo.predict(d)
7     predictor[['ds', 'yhat', 'yhat_lower', 'yhat_upper']].tail()
8     return predictor
9
10

```

Figura 27: Función para usar el modelo (Elaboración propia)

5. Evaluación de los resultados

Una vez entrenado el modelo, se realizan las pruebas con un rango de fechas

```

1 from datetime import datetime, timedelta
2 import math
3
4 inicio = '2022-06-01'
5 fin = '2022-06-05'
6
7 inicio = datetime.strptime(inicio, '%Y-%m-%d')
8 fin = datetime.strptime(fin, '%Y-%m-%d')
9
10 lista_fechas = [(inicio + timedelta(days=d)).strftime("%Y-%m-%d")
11                 for d in range((fin - inicio).days + 1)]
12 lista_cantidad=[]
13
14 for x in range(len(lista_fechas)) :
15
16     ventas_evaluar=prediccion_ventas(lista_fechas[x],modelo)
17     ventas_evaluar=math.ceil(abs(round(ventas_evaluar.yhat.values[0],2)))
18     print('Las ventas pronosticadas para el producto',numero_producto,' durante la fecha ', lista_fechas[x],' es de ', ventas_evaluar)
19     lista_cantidad.append(ventas_evaluar)
20
21
Las ventas pronosticadas para el producto 3 durante la fecha 2022-06-01 es de 1
Las ventas pronosticadas para el producto 3 durante la fecha 2022-06-02 es de 0
Las ventas pronosticadas para el producto 3 durante la fecha 2022-06-03 es de 0
Las ventas pronosticadas para el producto 3 durante la fecha 2022-06-04 es de 2
Las ventas pronosticadas para el producto 3 durante la fecha 2022-06-05 es de 2

```

Figura 28: Pruebas del modelo (Elaboración propia)

Para poder evaluar la precisión del modelo aplicamos la fórmula de Forecast Accuracy en el mes de Junio

$$\text{Fórmula 1: } 1 - \frac{|F - A|}{A}$$

F: Ventas predichas

A: Ventas reales

fecha	producto	Ventas	Ventas predichas
01/06/2022	3	0	1
02/06/2022	3	0	0
03/06/2022	3	4	0
04/06/2022	3	0	2
05/06/2022	3	0	2
06/06/2022	3	4	2
07/06/2022	3	7	3
08/06/2022	3	13	1
09/06/2022	3	0	1
10/06/2022	3	3	1
11/06/2022	3	0	2
12/06/2022	3	0	2

13/06/2022	3	0	2
14/06/2022	3	0	3
15/06/2022	3	1	1
16/06/2022	3	4	1
17/06/2022	3	0	1
18/06/2022	3	2	1
19/06/2022	3	0	2
20/06/2022	3	0	2
21/06/2022	3	5	3
22/06/2022	3	9	1
23/06/2022	3	0	1
24/06/2022	3	0	1
25/06/2022	3	4	1
26/06/2022	3	0	2
27/06/2022	3	0	1
28/06/2022	3	0	3
29/06/2022	3	0	1
30/06/2022	3	0	1
		Suma ventas	Suma ventas predichas
		56	45

$$FA = \left(1 - \frac{|45-56|}{56}\right) * 100 = 80.36\%$$

6. Implementación

Se procede a guardar el modelo en un archivo externo para poder usarlo en cualquier cuaderno de Colab.

```

1 # Aqui se guarda en la direccion que aparece en la ruta
2 from prophet.serialize import model_to_json, model_from_json
3
4 with open('/content/sample_data/serialized_model.json', 'w') as fout:
5     fout.write(model_to_json(modelo)) # Save model
6
7

```

Figura 29: Exportación del modelo (Elaboración propia)

ANEXO N° 20: Manual de usuario

Pantalla de ingreso a la aplicación que consta de dos apartados filtros del modelo y ejecución del modelo.

1. Filtros del modelo
2. Procesamiento del modelo

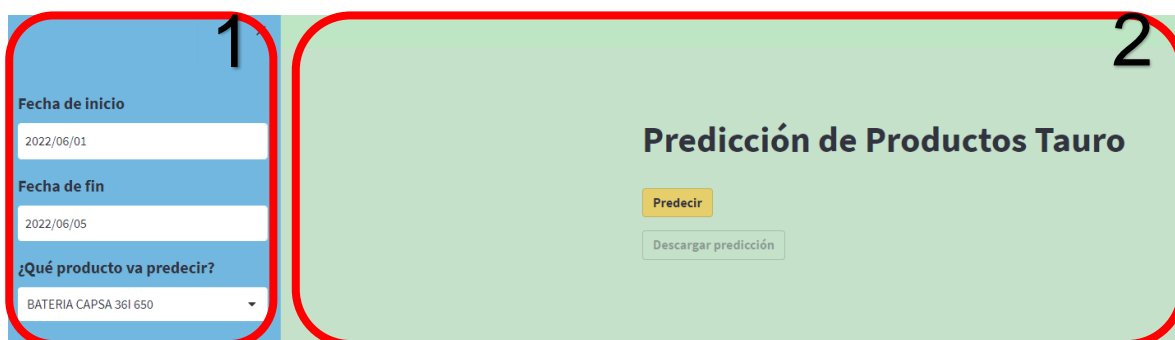


Figura 30: Pantalla de inicio del aplicativo (Elaboración propia)

En la siguiente imagen se muestran los filtros se van a elegir para mostrar la predicción.

Fecha de inicio: es la fecha desde donde se va realizar la predicción.

Fecha de fin: es la fecha donde finaliza la predicción.

Producto a elegir: es el producto que se va predecir.

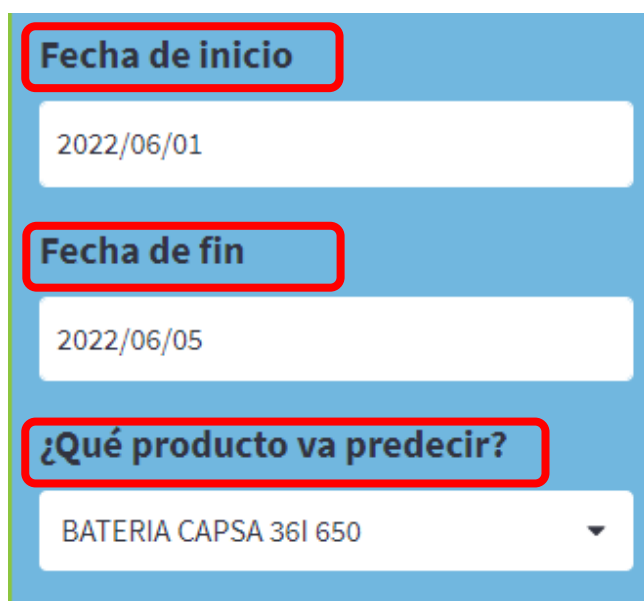


Figura 31: Filtros de la predicción (Elaboración propia)

En la siguiente imagen se muestra el botón para realizar la predicción. Cuando se presiona el botón “Predecir”, nos va mostrar un formulario por producto, fecha y cantidad predicha de acuerdo a los filtros especificados.



Figura 32: Botón para realizar la predicción (Elaboración propia)

Además de mostrar el formulario se va habilitar el botón “Descargar predicción”, donde al presionarlo se van a exportar los datos predichos en un archivo csv con nombre “Predicción”.



Figura 33: Botón para descargar la predicción (Elaboración propia)

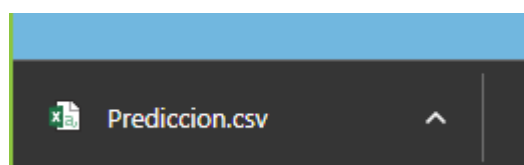


Figura 34: Archivo con la predicción (Elaboración propia)



UNIVERSIDAD CÉSAR VALLEJO

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

Declaratoria de Autenticidad del Asesor

Yo, FERNANDO MENDOZA APAZA, docente de la FACULTAD DE INGENIERÍA Y ARQUITECTURA de la escuela profesional de INGENIERÍA DE SISTEMAS de la UNIVERSIDAD CÉSAR VALLEJO SAC - LIMA NORTE, asesor de Tesis Completa titulada: "SISTEMA DE ANÁLISIS PREDICTIVO PARA LA GESTIÓN DE VENTAS DE LA EMPRESA TAURO A&F SAC", cuyos autores son ARAGON FARFAN CHRISTIAN JESUS JUNIOR, SANCHEZ ALTAMIRANO CLEIVER DERMALI, constato que la investigación tiene un índice de similitud de 23.00%, verificable en el reporte de originalidad del programa Turnitin, el cual ha sido realizado sin filtros, ni exclusiones.

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

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

LIMA, 22 de Diciembre del 2022

Apellidos y Nombres del Asesor:	Firma
FERNANDO MENDOZA APAZA DNI: 10363032 ORCID: 0000-0001-7981-8291	Firmado electrónicamente por: FEMENDOZAAPA el 27-12-2022 22:50:59

Código documento Trilce: TRI - 0499004