



Universidad
Señor de Sipán

**FACULTAD DE INGENIERÍA, ARQUITECTURA Y
URBANISMO**

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS

TESIS

**Modelo predictivo del proceso de ventas utilizando
inteligencia de negocios y data analytics en la
empresa centro textil De la Matta S.A.C.**

**PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO
DE SISTEMAS**

Autor:

Bach. Carreño Guerrero, Santiago Aníbal

ORCID: <https://orcid.org/0000-0002-4317-3687>

Asesor:

Mg. Chirinos Mundaca, Carlos Alberto

ORCID: <https://orcid.org/0000-0002-6733-8992>

Línea de Investigación:

Infraestructura, Tecnología y Medio Ambiente

Pimentel - Perú

2023

**MODELO PREDICTIVO DEL PROCESO DE VENTAS UTILIZANDO
INTELIGENCIA DE NEGOCIOS Y DATA ANALITICS EN LA EMPRESA
CENTRO TEXTIL DE LA MATTA S.A.C**

Aprobación del jurado

**MG. BRAVO RUIZ JAIME ARTURO
Presidente del Jurado de Tesis**

**MG. MEJIA CABRERA HEBER IVAN
Secretario del Jurado de Tesis**

**Dr. TUESTA MONTEZA VICTOR ALEXCI
Vocal del Jurado de Tesis**



Universidad
Señor de Sipán

DECLARACION JURADA DE ORIGINALIDAD

Quien suscribe la **DECLARACION JURADA**, soy Santiago Aníbal Carreño Guerrero del Programa de Estudios de Ingeniería de Sistemas de la Universidad Señor de Sipán S.A.C, declaro bajo juramento que soy autor del trabajo titulado:

MODELO PREDICTIVO DEL PROCESO DE VENTAS UTILIZANDO INTELIGENCIA DE NEGOCIOS Y DATA ANALITICS EN LA EMPRESA CENTRO TEXTIL DE LA MATTA S.A.C

El texto de mi trabajo de investigación responde y respeta lo indicado en el Código de Ética del Comité Institucional de Ética en Investigación de la Universidad Señor de Sipán (CIEI USS) conforme a los principios y lineamientos detallados en dicho documento, en relación con las citas y referencias bibliográficas, respetando al derecho de propiedad intelectual, por lo cual informo que la investigación cumple con ser inédito, original y autentico.

En virtud de lo antes mencionado, firma:

Carreño Guerrero Santiago Aníbal	06813643		
----------------------------------	----------	---	---

Pimentel, 15 de marzo de 2023

DEDICATORIA

Este trabajo de Tesis se lo dedico a Dios, que me sigue brindando la fuerzas y la vida para lograr mis metas y mis sueños; a mis padres, a mi hija a mi esposa, a mis hermanos, a mi familia, a mis amigos y todas aquellas personas que me dieron su apoyo incondicional sacrificando su tiempo y dándome fuerza para poder culminar esta carrera; a la Universidad y a todos los docentes que profesionalmente hicieron de mí una persona preparada y con disposición de seguir asumiendo retos en el transcurso de mi vida profesional.

Santiago Aníbal Carreño Guerrero

AGRADECIMIENTOS

A Dios por brindarme sabiduría y sus enseñanzas y que me ha permitido seguir cada día adelante, forjándome como un gran profesional.

A mis padres, Nolberto Aníbal y Luz Aurora por sus enseñanzas en sus valores y a seguir adelante sin desvanecer

A mi hija Jocelyn por su apoyo y por permitirme ser su ejemplo como profesional, cumpliendo mi sueño y sus sueños

A mi esposa, Verónica por su apoyo incondicional, su sacrificio y su confianza de todos los días

A mis hermanos, Luis, Yenny, Wilson, Geerson, Tania a pesar de la distancia brindándome su empuje cada día.

A mis amigos y amigas, por brindarme su apoyo, sus consejos y su tiempo

A todos los docentes que mediante sus enseñanzas me permiten ser un gran profesional.

A nuestro Profesor Ing. Heber Iván Mejía Cabrera

Al Ing. Luis Paredes Rosario

Santiago Aníbal Carreño Guerrero

RESUMEN

En el Perú, más específicamente, en la ciudad capital, Lima, distrito de La Victoria, se encuentra ubicado el Emporio Comercial Gamarra, donde existen una gran cantidad de empresas del sector textil, que se dedican a la comercialización de telas y prendas a nivel nacional; la gran mayoría utilizan sistemas automatizados para sus procesos de ventas, sin embargo, no ofrecen información para el pronóstico de ventas. La empresa Centro Textil de la Matta S.A.C., localizada en dicho emporio, se dedica a la actividad de producción y comercialización de telas deportivas a nivel internacional y nacional. Actualmente cuenta con un sistema automatizado para su proceso de ventas, sin embargo, éste no le permite obtener la información para pronosticar las ventas futuras, con lo que podría analizar e implantar estrategias de negocio, para lograr la planificación de la producción y la estabilidad financiera de la empresa. En este sentido con el propósito de lograr el mejoramiento de la toma de decisiones sobre las ventas y evaluar las estrategias de las mismas, se construyó un modelo de pronóstico de ventas utilizando redes neuronales, aplicando la técnica de Machine Learning usando Business Intelligence y Data Analytics que permitan pronosticar las ventas futuras. En el modelo se consideró un dataset de 1,460 registros de ventas históricas, que pertenecieron a las ventas diarias de los años 2017, 2018, 2019, 2020 y 2021, información con la que se realiza las proyecciones de ventas; utilizando la red neuronal de BP (Back Propagation) para pronosticar las ventas semanales (7 días). Asimismo, Asimismo, se utilizó la suite Anaconda Navigator, junto a la herramienta de programación Python y spider como editor. Obteniendo como resultado del error cuadrático medio (MSE) de 0.1358 y error absoluto medio (MAE) de 0.3110, demostrando que el modelo es eficiente para predecir las ventas. Se concluyó que, con el desarrollo de un modelo predictivo utilizando inteligencia de negocios y data analytics, sí se logrará la predicción de ventas en la empresa Centro Textil de la Matta S.A.C.

Palabras Clave: Proceso de ventas, predicción, aprendizaje automático, Backpropagation, inteligencia de negocios, data analytics.

ABSTRACT

In Peru, more specifically, in the capital city, Lima, district of La Victoria, is located the Emporio Comercial Gamarra, where there are a large number of companies in the textile sector, which are engaged in the marketing of fabrics and garments nationwide; the vast majority use automated systems for their sales processes, however, they do not provide information for sales forecasting. Centro Textil de la Matta S.A.C., located in this emporium, is engaged in the production and marketing of sports fabrics at the national and international level. It currently has an automated system for its sales process, however, this does not allow it to obtain the information to forecast future sales, with which it could analyze and implement business strategies to achieve production planning and financial stability of the company. In this sense, with the purpose of improving sales decision making and evaluating sales strategies, a sales forecasting model was built using neural networks, applying the Machine Learning technique using Business Intelligence and Data Analytics to forecast future sales. In the model, a dataset of 1,460 historical sales records was considered, which belonged to the daily sales of the years 2017, 2018, 2019, 2019, 2020 and 2021, information with which the sales projections are made; using the BP neural network (Back Propagation) to forecast weekly sales (7 days). Likewise, the Anaconda Navigator suite was used, together with the Python programming tool and spider as editor. As a result, the mean square error (MSE) of 0.1358 and mean absolute error (MAE) of 0.3110 were obtained, demonstrating that the model is efficient in predicting sales. It was concluded that, with the development of a predictive model using business intelligence and data analytics, the prediction of sales in the company Centro Textil de la Matta S.A.C. will be achieved.

Keywords: Sales process, prediction, machine learning, Backpropagation, business intelligence, data analytics.

ÍNDICE

I.	INTRODUCCIÓN	16
1.1.	Realidad Problemática	16
1.2.	Trabajos previos.....	26
1.3.	Teorías relacionadas al tema	39
1.4.	Formulación del Problema	75
1.5.	Justificación e importancia del estudio	75
1.6.	Hipótesis	75
1.7.	Objetivos	75
1.7.1.	Objetivo general.....	75
1.7.2.	Objetivos específicos.....	76
II.	MATERIAL Y MÉTODO	77
2.1.	Tipo y Diseño de Investigación	77
2.2.	Población y muestra.....	77
2.3.	Variables, Operacionalización	79
2.4.	Técnicas e instrumentos de recolección de datos, validez y confiabilidad 81	
2.5.	Procedimiento de análisis de datos.....	84
2.6.	Criterios éticos	87
2.7.	Criterios de Rigor Científico	88
III.	RESULTADOS	89
3.1.	Resultados en tablas y figuras	89
3.2.	Discusión de resultados	94
3.3.	Aporte práctico	97
IV.	CONCLUSIONES Y RECOMENDACIONES.....	178

4.1. Conclusiones.....	178
4.2. Recomendaciones.....	179
REFERENCIAS.....	180
ANEXOS	189

ÍNDICE DE FIGURAS

Figura 1. <i>Resumen del mercado mundial de Business Intelligence y Data Analytics 2020-2027</i>	18
Figura 2. <i>Mercado mundial de Business Intelligence y Data Analytics 2020-2027 por sectores</i>	19
Figura 3. <i>Perú: Empresas manufactureras según actividad económica</i>	20
Figura 4. <i>Lima: Microempresas manufactureras según actividad económica</i>	21
Figura 5. <i>Madurez digital por sectores económicos en el Perú</i>	22
Figura 6. <i>Inversión anual en innovación respecto al total de ventas durante el periodo 2015 -2017</i>	23
Figura 7. <i>Principales obstáculos para ejecutar actividades de innovación, 2015 - 2017</i>	24
Figura 8. <i>Información y las funciones gerenciales</i>	40
Figura 9. <i>Técnicas y herramientas de BI</i>	43
Figura 10. <i>Proceso de analítica predictiva</i>	52
Figura 11. <i>Clasificación de la analítica</i>	54
Figura 12. <i>Etapas del Método Delphi</i>	54
Figura 13. <i>Modelo de Aplicación ARIMA</i>	57
Figura 14. <i>Fases del Modelo CRISP-DM</i>	60
Figura 15. <i>Tareas y subtareas del proceso CRISP-DM</i>	62
Figura 16. <i>Aplicaciones del machine Learning</i>	69
Figura 17. <i>Modelo de red neuronal</i>	70
Figura 18. <i>Red neuronal Feedforward</i>	70
Figura 19. <i>Optimizador que implementa el algoritmo de Adam</i>	72
Figura 20. <i>Optimizador que implementa el algoritmo Adadelta</i>	73

Figura 21. Optimizador que implementa el algoritmo Adagrad.....	73
Figura 22. Optimizador que implementa el algoritmo Adamax.....	74
Figura 23. Optimizador que implementa el algoritmo RMSprop.....	75
Figura 24. Método para el desarrollo del informe de investigación	97
Figura 25. Pasos del método de revisión de la literatura utilizado.....	100
Figura 26. Resultado del proceso de selección.....	104
Figura 27. Modelos utilizados en la predicción de ventas	108
Figura 28. Representación gráfica del modelo predictivo propuesto.....	109
Figura 29. Información histórica de las ventas utilizadas para el proceso de ingreso del modelo predictivo propuesto.....	112
Figura 30. Información histórica de las ventas realizadas en los meses de agosto y septiembre 2020-2021.....	113
Figura 31. Entrenamiento para el proceso de análisis del modelo predictivo propuesto	115
Figura 32. Predicción para el proceso de análisis del modelo predictivo propuesto	118
Figura 33. Comparación de datos históricos vs datos predictivos.....	119
Figura 34. Esquema de la Red Neuronal para la solución del pronóstico de las ventas.....	121
Figura 35. Script de la arquitectura de la red neuronal en Python.....	121
Figura 36. Modelo de base datos DB SQL Matta.....	123
Figura 37. Stored procedure para transacción BD SQL Server a Excel.....	123
Figura 38. Datos obtenidos posterior a la limpieza de datos.....	126
Figura 39. Proceso de Inicio en Python (importación de librerías a utilizar).....	127
Figura 40. Script cargar de datos del dataset en python	127
Figura 41. Proceso de cálculo de valores de venta por día acumulado	128

Figura 42. <i>Script conversión de pesos de ventas de los 7 días de predicción en Python</i>	151
Figura 43. <i>Ejecución de procedimiento de set entrenamiento y validación</i>	154
Figura 44. <i>Ejecución del SCRIP de la red feedforward utilizando el optimizador ADAM</i>	155
Figura 45. <i>Ejecución del SCRIP entrenamiento de la red 70 épocas</i>	156
Figura 46. <i>Ejecución del SCRIP para visualizar los resultados de entrenamiento de la red 70 épocas</i>	156
Figura 47. <i>Ejecución del SCRIP para visualizar los resultados de pronóstico</i> ...	161
Figura 48. <i>Ejecución del SCRIP para visualizar los datos reales de los días 28/08/2021 al 23/09/2021</i>	165
Figura 49. <i>Valores reales de los últimos días del mes de setiembre del año 2021</i>	166
Figura 50. <i>Ejecución del SCRIP para visualizar los pesos correspondientes de los días 16/09/2021 al 30/09/2021</i>	167
Figura 51. <i>Valores de la predicción de 7 días</i>	168
Figura 52. <i>Diagrama de dispersión de 70 épocas</i>	171
Figura 53. <i>Diagrama de dispersión de validate loss</i>	172
Figura 54. <i>Cuadro estadístico del pronóstico de ventas de los últimos 22 días para el test de entrenamiento del modelo</i>	175
Figura 55. <i>Cuadro estadístico del pronóstico de ventas de los 7 días</i>	177

ÍNDICE DE TABLAS

Tabla 1. Características de las RNA	64
Tabla 2. Población.....	78
Tabla 3. Muestra	79
Tabla 4. Operacionalización de variables.....	80
Tabla 5. Resumen de campos de las fichas de recolección.....	82
Tabla 6. Resumen de fichas de registro electrónico.....	83
Tabla 7. Resumen de guías de entrevista	84
Tabla 8. Resumen resultante de los pesos de los 29 días	93
Tabla 9. Estudio acerca de modelos de aprendizaje automático	99
Tabla 10. Estructura del Método PICOC	101
Tabla 11. Estructura PICOC de la investigación	101
Tabla 12. Preguntas de investigación	102
Tabla 13. Fuentes de información empleadas.....	103
Tabla 14. Criterios de inclusión y exclusión	106
Tabla 15. <i>Papers obtenidos de la revisión bibliográfica</i>	107
Tabla 16. Comparación del desempeño de los modelos encontrados	108
Tabla 17. Variables existentes en las ventas históricas	124
Tabla 18. Ranking de valores para selección de variables a utilizar	125
Tabla 19. Datos de los años 2017-2018-2019-2020-2021 de valores para selección de variables a utilizar.....	129
Tabla 20. Criterios de inclusión y exclusión para limpieza de datos.....	147
Tabla 21. Datos de Valores nulos y extraños	148
Tabla 22. Datos para el test de validación	153

Tabla 23. Resultados de entrenamiento de las 70 épocas.....	157
Tabla 24. Resultados de pronósticos del test de validación del dataset	161
Tabla 25. Valores del test de validación de los pesos asignados en los valores reales y resultados predictivos	163
Tabla 26. Valores reales de los últimos días del mes de setiembre del año 2021	165
Tabla 27. Valores de la predicción de 7 días.....	167
Tabla 29. Valores de la Perdida del test entrenamiento	169
Tabla 30. Valores de los datos óptimos por año del data set.....	173
Tabla 30. Valores de los datos predictivos de los últimos 22 días del test de validación	174
Tabla 31. Valores de los datos de los 7 días pronosticados	176

ÍNDICE DE ANEXOS

Anexo 1. Resolución de aprobación del proyecto de investigación.....	189
Anexo 2. Carta de aceptación de la institución para la recolección de datos	192
Anexo 3. Instrumentos de recolección de datos - Variable Independiente.....	193
Anexo 4. Instrumentos de recolección de datos - Variable Dependiente	196
Anexo 5. Instrumentos de recolección de datos previos	199
Anexo 6. Data de ventas de la empresa Centro Textil de la Matta SAC, 2017 ..	202
Anexo 7. Data de ventas de la empresa Centro Textil de la Matta SAC, 2018 ..	203
Anexo 8. Data de ventas de la empresa Centro Textil de la Matta SAC, 2019 ..	204
Anexo 9. Data de ventas de la empresa Centro Textil de la Matta SAC, 2020 ..	205
Anexo 10. Data de ventas de la empresa Centro Textil de la Matta SAC, 2021	206
Anexo 11. Resultados del indicador “Tiempo total del proceso”.....	207
Anexo 12. Resultados del indicador “Memoria consumida durante un proceso”	209
Anexo 13. Resultados del indicador “CPU consumido durante un proceso”	210
Anexo 14. Resultados de entrevistas a personal de Centro Textil de la Matta ..	212
Anexo 15. Evidencias fotográficas.	215

I. INTRODUCCIÓN

1.1. Realidad Problemática

En la actualidad, en un contexto de negocios que cada vez es más desafiante y competitivo, las empresas de los diversos rubros económicos requieren de soluciones que les permitan generar ventajas competitivas por intermedio de “la recopilación, análisis y transformación de datos en decisiones estratégicas que les permitan diseñar planes exitosos y gestionar adecuadamente las distintas áreas y departamentos que la componen” (Castellnou, 2022). Dentro de ese contexto es que surge la conceptualización de Inteligencia Empresarial (IE), la cual analiza los procesos de la empresa en tiempo pasado, presente y futuro, de modo que, conforman una base de conocimiento necesario que le permite a estas empresas tomar decisiones y acciones para la diversidad de sus operaciones, verbigracia, las ventas, pronósticos, etcétera.

Respecto a lo anterior, de acuerdo al tipo de análisis que se hagan de los datos, las empresas pueden hacer uso de estrategias para analizar datos históricos tales como, el Business Intelligence (BI por su acrónimo), o de estrategias para predecir escenarios futuros, tales como, el Data Analytics (DA por su acrónimo). Ambos conceptos se encuentran extremadamente conectados y entrelazados en su enfoque para resolver problemáticas empresariales, brindar información sobre datos pasados y presentes y definir decisiones futuras. Sin embargo, el BI se centra más en el momento presente de los datos, tomando la decisión en función de los conocimientos históricos, mientras que, el DA se centra en el modelado predictivo, las estadísticas avanzadas y el aprendizaje automático para evaluar lo que sucederá en el futuro.

Por una parte, el BI facilita la previsión de la información en el departamento de ventas de las empresas. Esto se logra utilizando los datos crudos, conocidos como datos raw, en tiempo real derivados de múltiples áreas de la empresa. El BI puede utilizar datos de ventas tanto actuales como históricos para proporcionar escenarios actuales precisos acerca de las ventas; facilitando análisis y decisiones basadas en datos (Rikhardsson & Yigitbasioglu, 2018). Los representantes de ventas elaboran sus estrategias y planes basándose en datos analizados de modo que, toman decisiones comerciales más informadas. Por

ejemplo, si ven que un área de ventas es débil, pondrán más esfuerzo en alcanzar los objetivos en esas áreas. Si esos análisis de datos son correctos, las empresas podrán utilizar estos conocimientos para implementar estrategias y procesos para lograr los objetivos organizacionales. Si los datos son incorrectos, se puede perder tiempo y esfuerzos valiosos en actividades improductivas (Alnoukari & Hanano, 2017).

Por otra parte, el DA facilita la predicción de posibles escenarios que puedan darse en cuanto a operaciones empresariales, por ejemplo, utilizando datos validos logran un pronóstico de ventas sólido que les permite a las empresas mejorar el rendimiento de estas. El DA utiliza inteligencia artificial y aprendizaje automático para recopilar datos, por ejemplo, acerca de clientes potenciales y clientes actuales y estudia sus comportamientos a lo largo del proceso de ventas. El DA aplica esos conocimientos a las previsiones de ventas y al rendimiento de la canalización. En otras palabras, el DA en ventas ayuda a las empresas a convertir una cantidad abrumadora de datos en información procesable clara y fácil de entender que permite anticiparse para alcanzar y conectarse mejor con futuros clientes, predecir tendencias comerciales y de mercado, a impulsar el proceso de establecimiento de objetivos de ventas, pronósticos de ventas con mayor precisión, etcétera (O'Connor, 2022).

En consideración con lo anterior, se puede identificar que existe diversidad de beneficios con el uso de ambos tipos de análisis. El BI, mediante la previsión de ventas hace que los datos complejos sean más fáciles de entender para los usuarios empresariales. La visualización de datos, o la transformación de datos en gráficos, tablas y representaciones visuales, puede hacer que incluso los pronósticos de producción y ventas más complejos sean más fáciles de analizar, interpretar y elaborar estrategias. Esto, a su vez, conduce a una mejor comprensión y una previsión simplificada (Kurnia, 2018). El DA engloba el proceso de recopilación, almacenamiento y análisis de datos de ventas, ayudando a las organizaciones de ventas a comprender mejor cómo se compara su desempeño con los objetivos generales de la empresa, ayuda a que los vendedores pueden dedicar menos tiempo a trabajar con malos prospectos,

ayuda a una mejor segmentación de sus clientes potenciales, maximizar el valor de vida del cliente (CLV), entre otros (O'Connor, 2022).

Como puede verse, ambos enfoques permiten mejorar los procesos de ventas, por lo que, a nivel mundial, las empresas de todos los sectores económicos vienen incrementando esfuerzos por contar con dichos enfoques para analizar sus datos históricos y predecir escenarios futuros en cuanto a sus ventas. Esto queda demostrado en el estudio realizado por la compañía canadiense Emergen Research (2021), la cual es una empresa de consultoría estratégica y estudios de mercado con énfasis en transformación digital. Este estudio determinó que, el mercado global de BI y DA para el año 2019 se había valorizado en \$27.24 mil millones y se espera que alcance para el año 2027 un valor de \$60.49 mil millones, ello por la creciente demanda de soluciones para la toma de decisiones basadas en datos en tiempo real en las diversas industrias. Las soluciones de BI y DA ayudan a definir y crear informes comerciales, paneles de toma de decisiones, generando una demanda creciente de este tipo de soluciones, por lo que se espera que se continúe impulsando el crecimiento de los mercados globales de BI y DA.

Figura 1

Resumen del mercado mundial de Business Intelligence y Data Analytics 2020-2027

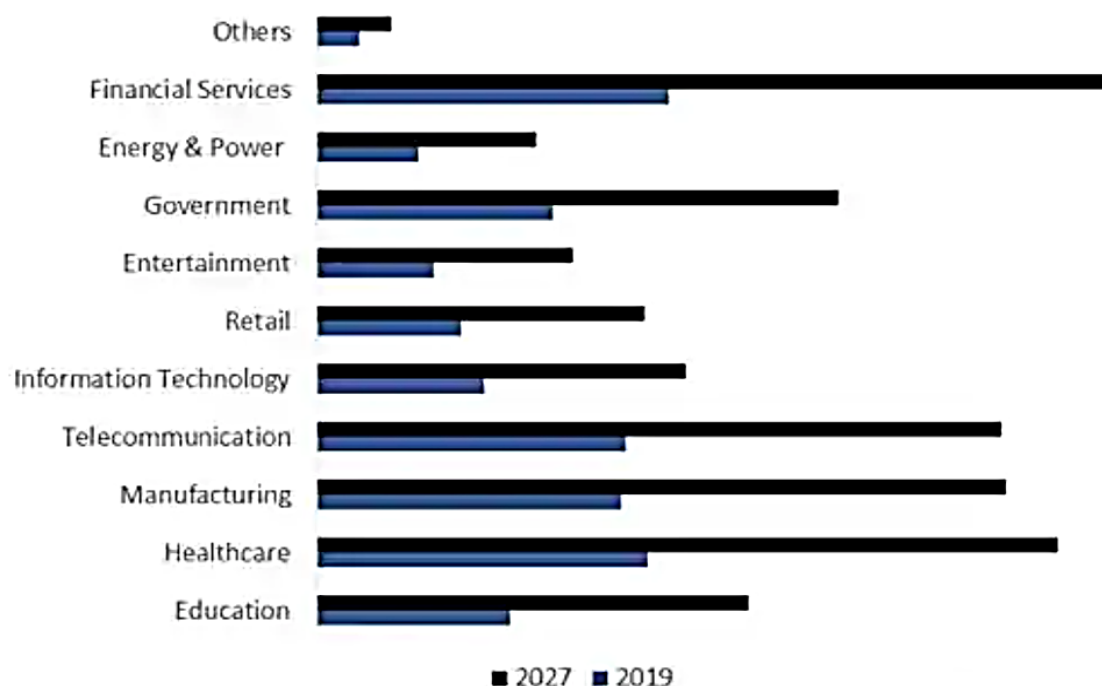


Nota. Fuente: Emergen Research (2021).

Dentro de ese mismo estudio, el mercado global de BI y DA está segmentado en educación, salud, manufactura, telecomunicaciones, tecnologías de la información, comercios minoristas, entretenimiento, gobierno, energía y electricidad, servicios financieros, etcétera. Se espera que el sector de servicios financieros lidere en términos de contribución de ingresos al mercado global de BI y DA para el periodo 2020-2027. Asimismo, se espera que los sectores salud, telecomunicaciones y manufactura aumenten sustancialmente ya que, las empresas de estos sectores han incrementado la demanda de soluciones analíticas para determinar las preferencias de sus clientes de manera más efectiva. Para ello, estas soluciones de BI y DA analizan los datos históricos de sus clientes, los productos comprados más recientemente y con mayor frecuencia, y preparan informes para identificar las tendencias de comportamiento de sus clientelas. Estas características están resultando muy beneficiosas y abren oportunidades importantes y más lucrativas para las empresas que operan en dichos sectores económicos.

Figura 2

Mercado mundial de Business Intelligence y Data Analytics 2020-2027 por sectores

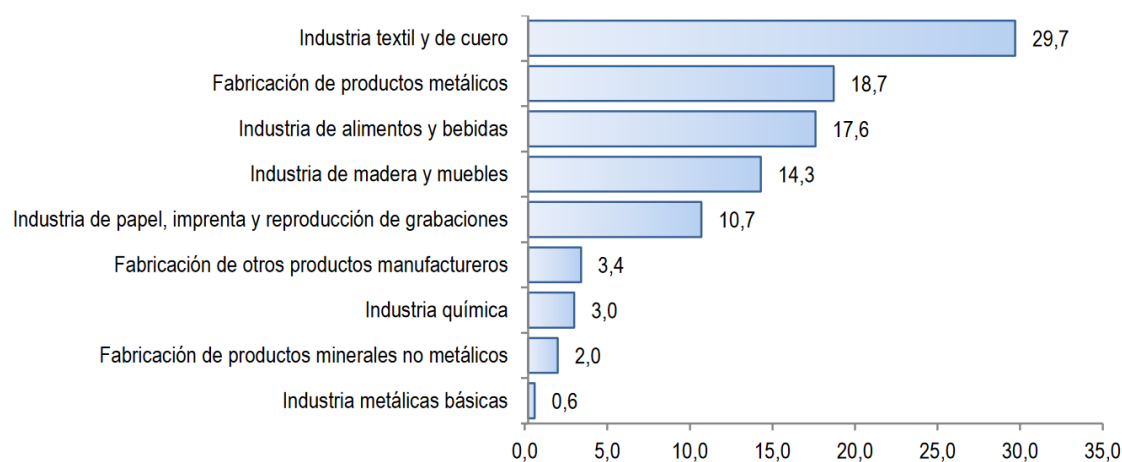


Nota. Fuente: Emergen Research (2021).

En el Perú, un estudio realizado por el Instituto Nacional de Estadística e Informática (en adelante INEI) definió a la actividad manufacturera como “la transformación física o química de la materia prima en productos destinados al consumidor, la cual es uno de los sectores económicos más importantes de la economía del país” (INEI, 2020, pág. 9). En dicho estudio, se identificó a la industria textil como actividad manufacturera más representativa (29.7%).

Figura 3

Perú: Empresas manufactureras según actividad económica



Nota. Fuente: INEI (2020).

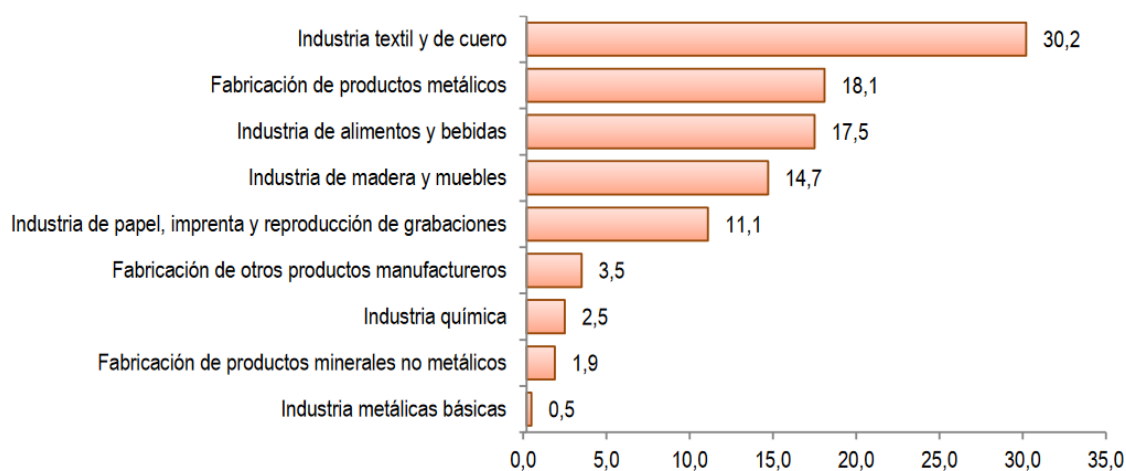
Por una parte, se evidenció que, el número de microempresas manufactureras ha disminuido para el año 2020 (199856 microempresas) en comparación con el año 2019 (198331 microempresas), teniendo una tasa de decrecimiento de 0.8%. Este sector se vio afectado por el decrecimiento de los niveles de ventas a causa de la pandemia por la COVID-19 (Moreno, 2020).

Por otra parte, del total de microempresas manufactureras del Perú, son 99895 las que se encuentran ubicadas en la Región Lima, destacando un total de 59884 empresas las que se dedican a la actividad económica de la industria textil y de cuero (30.2%). “Esto se explica por la gran concentración de unidades productivas de confección de prendas de vestir y que luego se ofrecen en los centros comerciales del país” (INEI, 2020, pág. 50). Esto se puede evidenciar

con el gran número de empresas de la actividad de industria textil y de cuero que se encuentran localizadas en el Emporio Comercial Gamarra.

Figura 4

Lima: Microempresas manufactureras según actividad económica



Nota. Fuente: INEI (2020).

Tal y como se mencionó anteriormente, la pandemia a causa de la COVID-19 aminoró los niveles de ventas de las empresas en todas las industrias, no siendo ajeno, el sector textil peruano, por lo que, en la era postpandemia, “el foco estará en la reactivación económica en dichas industrias, mediante la creación de soluciones para estas problemáticas sociales, por ejemplo, mediante el desarrollo de los modelos predictivos e investigaciones científicas...” (Universidad de Lima, 2021). Para lograr el desarrollo de dichos modelos predictivos, se tendrá que ejecutar el procesamiento de grandes volúmenes de datos apoyados en la inteligencia de negocios, el data analytics, la inteligencia artificial, el machine learning y entre otros recursos y otras herramientas más.

Respecto a lo anterior, un estudio realizado en el Perú por la empresa Ernst & Young (2022) reveló un retroceso en el índice de madurez digital en 2% lo que reflejó una tendencia de desaceleración coyuntural teniendo sectores muy afectados en la era postpandemia tal y como el sector manufactura, el cual también tuvo un retroceso porcentual llegando a los 57.81%.

Figura 5

Madurez digital por sectores económicos en el Perú



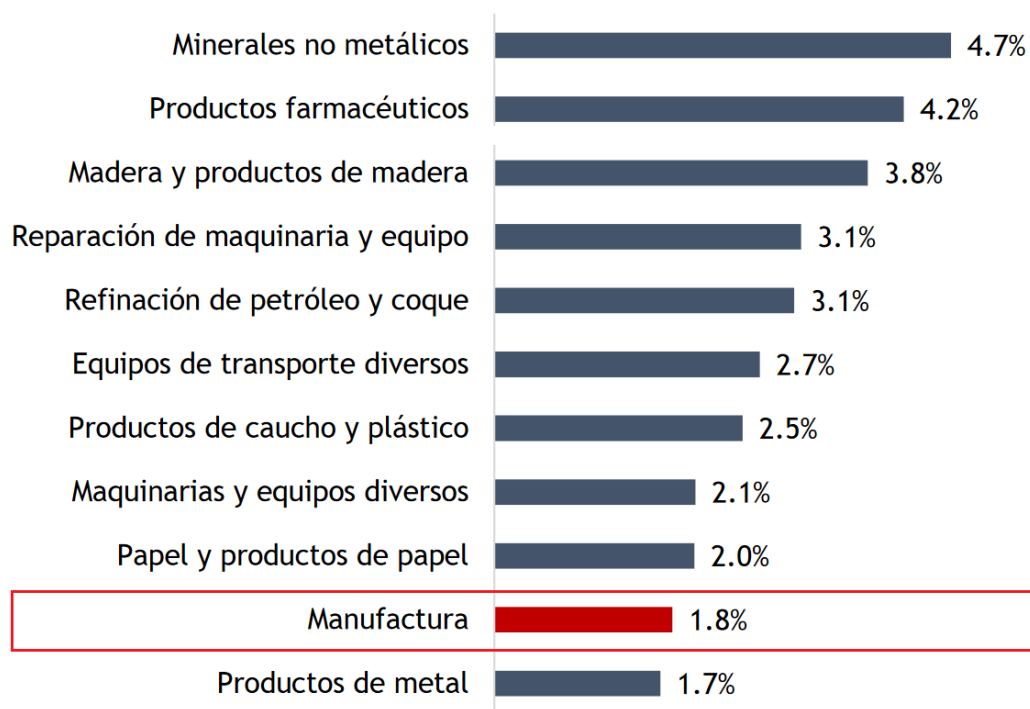
Nota. Fuente: Ernst & Young (2022).

Asimismo, este estudio reveló un incremento en cuanto a el presupuesto asignado para la transformación digital (6%) y teniendo como área prioritaria a las ventas y el marketing (31.97%). También es importante recalcar que, entre las tecnologías a utilizar para los procesos de transformación digital destacan aquellas asociadas a la inteligencia de negocios tales como, el big data y Analytics (51.02%). “Estas se han venido consolidando como la tecnología más utilizada en los últimos años para poder identificar tendencias de consumo entre los clientes y en el entorno actual donde los consumidores han venido adquiriendo bienes y servicios de manera digital, dejando una enorme huella de información en cada transacción” (Ernst & Young, 2022). Finalmente, se evidenció que, entre las barreras que las empresas peruanas destacan como principales se tienen a la carencia de personal capacitado (46.94%), falta de tiempo (34.69%), resistencia al cambio (34.01%), falta de presupuesto para costear los sistemas de inteligencia de negocios (27.21%), desinterés de los líderes (19.73%), entre otros.

Del mismo modo, un estudio realizado por el Ministerio de la Producción (2018) reveló que las empresas manufactureras peruanas son la segunda industria económica que menos invierte de manera más intensiva en innovación tecnológica con un 1.8% del total de ventas durante el periodo 2015-2017.

Figura 6

Inversión anual en innovación respecto al total de ventas durante el periodo 2015 -2017



Nota. Fuente: Ministerio de la Producción (2018).

Igualmente, se reveló que existen obstáculos que no permiten ejecutar actividades de innovación tecnológica en dicha industria manufacturera peruana:

Figura 7

Principales obstáculos para ejecutar actividades de innovación, 2015 -2017



Nota. Fuente: Ministerio de la Producción (2018).

Dentro de esas desalentadoras estadísticas del sector manufacturero es que se encuentran las empresas perteneciente al sector textil del Emporio Comercial Gamarra situado en el corazón del distrito de La Victoria, las cuales no innovan tecnológicamente para mejorar sus procesos, tales como el proceso de ventas, evidenciándose una brecha en cuanto al uso de un modelo predictivo que les permita proyectar sus futuras ventas, por lo que consideran que es muy costoso, que no brinda información que necesitan para satisfacer sus necesidades y tomar decisiones oportunas y acertadas acordes a sus necesidades, entre otros obstáculos. Por ello es que es necesario que estas empresas dispongan de un modelo predictivo del proceso de ventas el cual podría estar basado en la utilización de BI y DA.

Respecto a la problemática anterior, Calderón (2018) mencionó que, para llevar a cabo una buena toma de decisiones empresariales y que permitan el incremento de las ventas en las empresas textiles, es imprescindible contar con un buen planeamiento estratégico, para lo cual se hace necesario disponer del uso de herramientas de BI. Estas herramientas, permiten analizar datos históricos (por ejemplo, el BI) y permiten desarrollar modelos predictivos (por ejemplo, Data Analytics). Además, “cada vez más organizaciones peruanas

apuestan por las soluciones de BI ya que, permite agilizar procesos en todas las áreas y, en consecuencia, adquirir una ventaja competitiva en la industria” (ESAN, 2019).

Sin embargo, el problema principal respecto a lo mencionado anteriormente es que, en principio se tienen datos de diversas situaciones que se dan en las microempresas peruanas, por ejemplo, los datos reales de ventas históricas de las empresas textiles peruanas, pero para hacer DA o BI, se requieren de datos tengan consistencia, de data histórica válida, por lo que se hace trascendental la existencia de procesos ordenados en dichas organizaciones peruanas, por ello, se requiere de disciplina en el registro de tales datos, puesto que, de no registrarse estos adecuadamente, mostrarán inconsistencias y dicha data histórica ya no será válida, por ejemplo, suceden situaciones en las que, al momento de ingresar datos o información para la emisión de documentos (tales como facturas, boletas, notas de crédito, etcétera), existen datos de los que se han prescindido por error tal vez humano, los cuales posteriormente se graban y queda confirmado en los sistemas de información que emplean las microempresas peruanas. Por ello, son necesarios procesos de ventas ordenados, así como también, disciplina en el registro de datos respecto a estos procesos en este tipo de empresas textiles.

No obstante, en el Perú, las empresas manufactureras del sector textil son desordenadas e informales, evidenciándose problemas de inconsistencia y data histórica no válida en la ejecución de los procesos de ventas. Por tanto, en este tipo de escenario y con la necesidad de realizar DA, no pueden hacerse predicciones ya que dichas datas no son confiables y servibles al 100% ya que demuestran no ser válidas, no pudiéndose ejecutarse predicciones pues no se disponen de datas históricas que sirvan para dicho proceso predictivo. Por lo expuesto, queda evidenciado el problema de ingeniería en la siguiente interrogante: ¿cómo hacer DA de manera confiable con datos inconsistentes, con datos históricos no válidos en muchos aspectos, con datos de empresas que poseen procesos ampliamente desorganizados, generadoras de datas desordenadas, carentes de disciplina para el registro de datos? Por lo que, ese es el mayor problema, ¿cómo hacer DA en tal realidad? Por lo que se pretende

comprobar si las redes neuronales son lo suficientemente tolerantes para que puedan hacer DA, a pesar de estos escenarios no son ideales, sino que, por el contrario, son adversos para los procesos predictivos. Más específicamente, se hará uso de Inteligencia Artificial para hacer predicciones más adecuadas.

Un ejemplo de este tipo de empresas manufactureras del sector textil peruano que cuentan con problemas de datos históricos es la empresa caso de estudio, la empresa Centro Textil de la Matta SAC que se encuentra ubicada en dicho emporio comercial victoriano. Esta empresa utiliza un sistema automatizado para su proceso de ventas el cual no le permite obtener información sobre la proyección de ventas y tampoco cuenta con un área de análisis de datos, que le permita extraer y examinar dichos datos y así poder contar con los beneficios de llevar a cabo las estimaciones de ventas.

Por esta razón arriba mencionada, es imprescindible para la empresa disponer de un modelo predictivo para sus procesos de ventas el cual le permita obtener sus proyecciones estimadas y poder aplicar las estrategias necesarias de ventas, lo cual hace a las empresas “más competitivas en el mercado textil logrando tener una mejor cultura y gestión empresarial” (Álvarez, 2017).

Por esta razón, el presente informe se desplegó con el propósito de desarrollar un modelo predictivo del proceso de ventas utilizando BI y DA permitiéndole en la empresa Centro Textil de la Matta S.A.C., tener alta confiabilidad en los resultados de los pronósticos de las ventas en los días o meses posteriores, cumpliendo con esto con el objetivo estratégico del negocio el cual es el mercado textil.

1.2. Trabajos previos

Fan (2021), realizó la investigación, Prediction of A-share trend based on machine learning technologies, en China, Predicción de la tendencia A-share basada en la máquina. Caso de estudio en el uso de Las tecnologías de aprendizaje automático aplicado en el campo económico y financiero de china en la negociación de Acciones chinas de gran capitalización del CNY (la divisa

legal de la República Popular China). Sobre los inversores minoristas ordinarios que compran y venden una pequeña cantidad de acciones de bajo valor. Este estudio se centró principalmente en las acciones de gran capitalización en China. Los datos incluyen el ID de acciones de algunas acciones de gran capitalización, fecha y el tiempo, sus correspondientes características y si un stock específico puede aumentar un 5 por ciento en un tiempo específico. Para ello se realizó el procesamiento usando normalización Min-Max, estandarización y transformación Box-Cox, e implementación de cinco modelos de aprendizaje automático, incluido el modelo SVM, Naïve Modelo de Bayes, modelo de árbol de decisión Modelo de red neuronal BP y modelo Wide & Deep, respectivamente. Según el autor se basó en el uso de algoritmos de redes neuronales utilizando estrategias combinadas y óptimas para lograr los mejores resultados considerando que el aprendizaje máquina que al procesar los algoritmos en la transformación de datos apropiada nos brinda los errores correspondientes que permita acercarse con mayor precisión a la predicción por lo tanto, se asignó en el conjunto de datos el ID de las acciones y su características con la etiqueta "0" o "1" que representan si aumentan una tendencia del 5 por ciento en la transformación de la normalización. En el procedimiento de aprendizaje de la red neuronal BP se dividió principalmente en dos partes: la propagación hacia adelante de características de entrada y la propagación inversa del error entre producción y valor esperado el optimizador seleccionado fue el Adam, cambiando los pesos por la función de activación y la pérdida. Al implementarse cinco modelos diferentes en el stock pool y logró resultados satisfactorios, especialmente BP neural modelo de red obtuvo mejores resultados con la Precisión de 0,95. Mientras tanto, el AUC del modelo SVM alcanzó 0,881. Sin embargo, todavía existen algunas limitaciones, como la excelente El resultado solo puede obtenerse del conjunto de acciones específico y numerosas características detalladas. En conclusión, en este conjunto de acciones, Estos cinco métodos de aprendizaje automático son confiables en ayudar a los inversores minoristas en general a predecir acciones de gran capitalización.

Abdellatif et al. (2019), realizaron la investigación, Egyptian Case Study-Sales forecasting model for automotive section. Caso de estudio en una empresa Egipcia de un modelo de previsión de ventas para la sección de automoción realizado en una de las más grandes empresas de Egipto que fabrica automóviles, la investigación se basó en realizar el pronóstico de ventas en la cual juega un importante papel en las empresas de fabricantes de automóviles de los países a nivel mundial en este caso se utilizó como datos las variables más efectivas el volumen de ventas, como los precios de venta unitario, la tasa de inflación, el ingreso per cápita, el precio del combustible y también analizando los resultados de las ventas anteriores. Se utilizó el 15% de cada variable los datos mensuales de las ventas entre Feb-2015 hasta Mar-2018 como prueba. Los modelos utilizados fueron las redes neuronales, El Sistema de inferencia Neuro-difuso adaptativo (ANFIS) y Regresiones lineales múltiples (MLR) para ello se realizó la comparación de estos modelos y selecciono el mejor resultado. Obteniendo como resultado que la técnica ANN (Artificial Neural Networks) obtuvo el mejor resultado superior a las otras técnicas en siendo los valores de coeficientes, como en el valor R2 para el modelo ANN es igual a 0,67 y RMSE (El error cuadrático medio) es igual a 190,9, según sus datos propuestos lo que se llegó a obtener como resultado fue una gran capacidad de este modelo en la previsión de ventas de automóviles. Asimismo, la comparación de los modelos propuestos, muestran que los valores R2 y RMSE del modelo ANN son superiores a los del modelo ANFIS y MLR. y como resultado, la técnica ANN es supervisada en comparación con ANFIS y MLR y evaluada por RMSE y R2 en el pronóstico de ventas automotrices ((ANFIS Y MLR), como RMSE (626.79,276.59) Y R2 (0.30,0.29). Es así que, para los estudios futuros, la salida se utilizará como entrada para el componente de optimización para optimizar el plan de producción para lograr e implementar un modelo analítico prescriptivo el cual le permitirá cuantificar la toma de decisiones sobre los futuros procesos operacionales del plan de producción.

Wang et al. (2019), realizaron la investigación, Sales Forecast of Chinese New Energy Vehicles Based on Wavelet and BP Neural Network, en China, Previsión

de ventas de vehículos chinos de nueva energía basada en Wavelet y BP Neural Network. Durante las últimas décadas y, especialmente durante estos últimos años el mundo ha sufrido alteración en la utilización de nuevas energías para el uso de los vehículos teniendo como objetivo el cuidado del medio ambiente y para lo cual han ido ocupando cada vez mayores ventas en el mercado mundial. Para este estudio se utilizó una red neuronal para establecer un modelo de predicción. Analizando los factores de influencia que pueden afectar las ventas. el autor Su Yue hizo usanza de un modelo de regresión lineal y modelo gris para la predicción. La combinación de los dos modelos pudo reducir el error, pero tampoco se tuvo en cuenta la influencia de los factores que afectarían, una gran desviación en el resultado. Yan Fan predijo las ventas anuales de vehículos de nueva energía a través del Modelo logístico combinado con influencia política y progreso tecnológico. El modelo de manera integral consideró los dos factores, pero utilizando las ventas anuales y el volumen, y también obtuvo poco error y la precisión disminuyó. A partir de estos estudios, se pudo ver que si agregar más factores de influencia conducirán a una gran desviación en los resultados de predicción y ventas mensuales de los vehículos de nueva energía que son cíclicos. Por lo tanto, este documento toma las ventas mensuales como índice de análisis. En primer lugar, pronostica a través de datos de ventas de series de tiempo, y luego considera exhaustivamente los efectos de cuatro factores (factores estacionales, cambios en los precios del petróleo, nivel de desarrollo de nuevas tecnologías y políticas energéticas). para predecir las ventas. Donde consideró las ventas mensuales de enero de 2015 a junio de 2018 como muestra, y establece un modelo de predicción con el propósito de pronosticar las ventas de julio de 2018 a diciembre de 2018. En este artículo, se implementó el algoritmo de la red neuronal wavelet (RNW) donde se utiliza para predecir las ventas en función de series de tiempo. Siendo su precisión superior a la de las estadísticas generales. Al construir el modelo de red neuronal de pronóstico de ventas, la red neuronal BP (retro propagación) se estructuró en tres (03) capas, es decir, a) capa de entrada, b) capa oculta y c) capa de salida. considerando la capa de entrada como el volumen de ventas mensual de vehículos de enero de 2015 a junio de 2018, un total de 42 meses,

El nodo de la capa oculta se compone de la función wavelet, capa de producción de ventas mensuales de julio de 2018 a diciembre 2018. El resultado del error del mes de julio fue demasiado grande, porque el valor de la muestra en julio de 2016 y 2017 disminuyó en comparación con el mes de junio del año en curso, por lo que el valor previsto disminuye, mientras que el valor en julio de 2018 aumenta en la realidad. El error de predicción promedio de la red neuronal wavelet es 25,47%. Y se puede ver en los resultados que no es exacto. Por tanto, es necesario para mejorar el modelo y agregar factores de influencia que afectan las ventas en volumen, para el mejoramiento de la precisión de la predicción. NO es suficiente hacer predicciones solo a través de los datos de muestra de series de tiempo. Para hacer más precisos en la previsión de ventas, los factores de impacto deben cuantificarse y agregarlos en el modelo de predicción. Cuanto más fuerte sea la correlación entre el factor de influencia y las ventas de salida de volumen, mayor es la precisión de la predicción del modelo. Combinando estos factores con datos de muestra, se utilizó el modelo de predicción de la red neuronal de BP (Back Propagation) es una especie de feedforward multicapa red de acuerdo con el algoritmo de retro propagación logrando obtener mejores resultados de predicción más precisos que se ajusta según al cambio de factores influyentes en este tiempo. Obteniendo como resultado del error cuadrático medio (RMSE) de predicción es del 5,0%, lo que tiene un buen valor de aplicación determinada. Por tanto, la red neuronal de BP es el modelo de predicción que puede predecir con precisión el volumen de ventas de vehículos de nueva energía.

Punam et al. (2018), realizaron la investigación, A Two-Level Statistical Model for Big Mart Sales Prediction, en la India. Plantea Un modelo estadístico de dos niveles para la predicción de ventas de la empresa Big Mart. La previsión de ventas es un aspecto importante en los grandes mercados o empresas dedicadas al comercio minorista, logística, fabricación, comercialización y venta al por mayor. Esto les permite a las empresas asignar los recursos necesarios de manera eficiente, para estimar los ingresos por ventas alcanzables y planificar una mejor estrategia para crecimiento futuro de la empresa. En este

artículo se ha considerado la predicción de ventas para la empresa Big Mart que es una corporación minorista internacional, analizando un producto de venta en un punto de venta. Para lo cual el proceso predictivo se realizó en dos etapas la primera etapa se utilizó técnicas predictivas populares como lineal, regresión, árbol de regresión cubista regresión vectorial y k-vecinos más cercanos y la segunda etapa, donde se utilizó la estadística de dos niveles y se construyó el modelo estadístico de dos niveles que consistió en el aprendizaje automático como la regresión lineal regresión, vectorial y cubista logrando a combinar los algoritmos para lograr hacer la predicción final. Se aplicó el método de apilamiento que generalmente se usa para combinar las técnicas de aprendizaje automático para mejorar la precisión de los modelos predictivos. Se realizó el estudio de los datos de las ventas del año 2013, teniendo un conjunto de datos de 8523 observaciones y compuesta de 12 atributos. Esto se particionó en dos partes, datos de entrenamientos y datos de pruebas en una proporción de 80:20, es decir, 6818 instancias son datos de entrenamiento y el resto 1705 instancias datos de prueba. El resultado obtenido mediante el uso de un modelo estadístico de dos niveles fue el más preciso obteniendo la media absoluta con un valor de error hasta 39,17%. Siendo el modelo que superó las otras técnicas predictivas y contribuyó con mejores predicciones al conjunto de datos de Big Mart.

Bowen et al. (2020), realizaron la investigación, Forecasting method of e-commerce cargo sales based on ARIMA-BP model, en China. Método de previsión de las ventas de carga de comercio electrónico basado en el modelo ARIMA-BP en este artículo se analizado el comportamiento de los comerciantes sobre el flujo efectivo debido a un reabastecimiento irrazonable y el desperdicio de recursos y lucro cesante causados por bienes y escasez. La problemática del volumen de ventas del comercio electrónico es una complicación de predicción de series de tiempo, para obtener el período futuro de ventas de productos básicos, la satisfacción, la tasa de devolución y el volumen de ventas se consideran como datos de entrada para obtener los volúmenes de ventas de la comercialización electrónica en el estudio se ha utilizado dos modelos de

predicción separados, la red neuronal BP y el ARIMA, basándose en volumen de ventas de los próximos cinco días y estableciendo un modelo de error cuadrático medio para ponderar el ajuste y los resultados predictivos de las dos predicciones individuales. Donde el error en la red neuronal BP es 0,56415 y el error del modelo ARIMA es 1.67211 al realizar la comparación de ambos modelos se verifico que carecen de complementariedad y pueden extraer y analizar información de características de manera efectiva, para ello se propuso realizar un modelo de combinación no lineal con ponderación basado en la ecuación media de ARIMA-BP obteniendo como resultado el error normalizado de 0.05363. Siendo más estable la predicción combinada y el error es mucho menor que una sola predicción, Finalmente, los resultados del cálculo pueden probar la aplicabilidad del modelo de combinación ARIMA-BP. Según la investigación muestra que el modelo puede adaptarse al problema de pronosticar el volumen de ventas de bienes de comercio electrónico determinando que se puede aplicar bien a las ventas.

Zheng & Li (2018), realizaron la investigación, Unmanned Retails Distribution Strategy Based in Sales Forecasting, en Canadá. Estrategia de distribución de minoristas no tripulados basada en pronósticos de ventas. Con el creciente desarrollo de la tecnología móvil y las ventas con el nuevo concepto de "Nueva venta minorista", cada vez más capitalistas de riesgo han estado persiguiendo oportunidades de inversión en el negocio minorista no tripulado, como Future Store, Xiaomai, Xingbianli, CityBox, también y empresas de comercio electrónico. Alibaba Group Holdings y JD.com. En la actualidad las formas de venta minorista sin personal son las máquinas expendedoras inteligentes, y las tiendas de conveniencia sin personal y los estantes inteligentes sin personal en las oficinas. Estos formatos de venta tienen una gran ventaja: con la ayuda de las nuevas tecnologías, el costo del comercio minorista no tripulado es mucho menor que el de los formatos de minoristas tradicionales. Sin embargo, la forma de formular una estrategia eficaz de reabastecimiento y distribución es el mayor problema operativo, también es la principal pesquisa investigativa de dicho artículo. En primer lugar, se obtiene los datos históricos de ventas de los nodos

a través de los registros de pago en línea y luego se pronostica las ventas futuras utilizando métodos de pronóstico más comunes que son la media móvil, el suavizado exponencial y el modelo de media móvil autorregresiva (ARMA). Se realizó una prueba con la utilización de 20 nodos usando los registros de ventas de los diez días anteriores para pronosticar la demanda del día se desarrolló el método de Prueba de Raíz Unitaria dando como resultado 0.01 y luego se realizó la optimización de las rutas de distribución para poder reabastecer todos los nodos agotados para ello se consideró el peso de carga que puede trasladar cada vehículo y cuantos nodos podría trasladar de forma se optimiza las rutas distributivas para aminorar efectivamente los costes de transporte .De todos los modelos el que funciona mejor en la predicción a corto plazo de series de tiempo. Es el modelo ARMA (media móvil autorregresivo) el cual expresa el valor de secuencia como una suma ponderada de valores pasados y elementos de perturbación pasados para pronosticar las demandas de los nodos en conclusión del estudio se predetermina principalmente dos problemas actuales en las ventas no tripuladas: la previsión de ventas y la optimización de la distribución obteniendo el esquema de distribución óptimo. Debido a la situación real y la complejidad de la investigación, los modelos de este artículo se simplificaron. Por ejemplo, este artículo solo estudió el mismo tipo de producto, mientras que en la operación real generalmente hay muchas categorías diferentes juntas, por lo que es necesario seguir estudiando y mejorando.

Dairu & Shilong (2021), realizaron la investigación, Machine Learning Model for Sales Forecasting by Using XGBoost, en China. Modelo de aprendizaje automático para la previsión de ventas por Usando XGBoost. Las corporaciones minoristas modernas que operan enorme cadena de negocios, la predicción de ventas exacta es la clave para impulsar el desarrollo de las corporaciones, incluso pueden establecer el fracaso o éxito de la misma. Por esta razón, se propone un modelo de predicción de ventas preciso mediante el aprendizaje automático. Tomando como ejemplo la información a la empresa Walmart, Inicialmente la ingeniería extrae ciertas características de los datos

históricos de ventas. Para poder hacer aplicados a los modelos eXtreme Gradient Boosting (XGBoost) para pronosticar el monto de las ventas futuras, asumen como objetivo diseñar un algoritmo de pronóstico fundamentado en XGBoost para lograr mejores rendimientos. Para ser más específicos, se determinó dos áreas, la clasificación y regresión. Para construir el conocimiento de las ventas de mercado a partir de datos estadísticos de ventas durante los últimos 1913 días de las tiendas Walmart en tres estados y pronosticar las ventas en los próximos 28 días. En comparación con otros métodos de aprendizaje en conjunto, XGBoost se ejecuta diez veces más rápido utilizando mucho menos recursos, es sensible a los valores atípicos, primero hacemos un preprocesamiento de datos para filtrar valores atípicos, luego se convierten en tipos de datos flotantes para guardar en memoria. Luego, la característica de tiempo agregado se extrae de diferentes períodos de tiempo de datos de ventas de día, semana y mes a año. Para evaluar de forma eficaz nuestro algoritmo, se realizó varios experimentos utilizando Kaggle donde existe un conjunto de datos de competencia. Que contiene datos recopilados de 1913 días de la corporación Walmart. Los resultados experimentales demuestran que nuestro algoritmo brinda resultados de vanguardia. La puntuación RMSSE en los modelos Para RMSSE (error cuadrático medio de la velocidad) el puntaje, en el modelo XGBoost obtiene el más bajo 0.655, mientras que el modelo clásico de regresión lineal obtiene 0,783, 19,5% más alto que el modelo XGBoost. y el modelo de regresión de cresta obtenga 0.774, también 13.6%. Los resultados experimentales demuestran que el modelo propuesto basado en XGBoost tiene un buen rendimiento para mejorar la velocidad de ejecución y precisión de la predicción. Mientras tanto, es posible explorar más características eficientes en la ingeniería de características para mejorar aún más pronóstico del desempeño.

Dong et al. (2021), realizaron la investigación, A Machine Learning Model for Product Fraud Detection Based On SVM, en China, Una metodología de aprendizaje automático con el propósito de lograr la detección de fraudes de productos basado en SVM. Con el auge de la tecnología internet de las cosas

cada vez más. Las empresas utilizan esta tecnología para la producción del trabajo diario. Esta tecnología generará grandes cantidades de datos durante el proceso de solicitud. Si los datos se pueden usar con prudencia, ayudarán las empresas toman mejores decisiones. En este artículo, proponemos una metodología de predicción de fraudes fundamentada en el Modelo de clasificación SVM (algoritmo de clasificación y regresión). En la cadena de suministro de muchas industrias, existe información incompleta de compradores y vendedores, operaciones irregulares y empresas que no son accesibles al crédito a menudo se producen fraudes, incumplimiento a tiempo y riesgos crediticios. Las tecnologías de aprendizaje automático y minería de datos pueden analizar datos previos de la cadena de suministro para determinar si existe fraude en la circulación de productos. Se realizará con un conjunto de datos perteneciente a la empresa DataGo. Para análisis y preprocesamiento de datos, se utilizarán tres (03) algoritmos de Naive-Bayes, Árbol de decisión y máquinas de vectores de soporte, primero realizamos la ingeniería de características de datos para que estos puedan ser utilizados por el modelo de clasificación SVM el cual es de aprendizaje supervisado para ello se consideró datos como el suministro, producción, ventas, etc. Los experimentos demostraron las siguientes precisiones Regresión logística 97,80, Bayes ingenuo de Gaus 87,84, y SVM 98.61 Finalmente, encontramos que el modelo basado en SVM es el de mejor capacidad de predicción en la clasificación de fraudes de productos básicos.

Pavlyshenko (2019), realizó la investigación, Machine-Learning Models for Sales Time Series Forecasting, en Ucrania, Modelos de aprendizaje automático para el tiempo de ventas Predicción de series. En este artículo, estudiamos la usanza de modelos de aprendizaje automático para el análisis predictivo de ventas. El objetivo principal es considerar los principales enfoques y estudios de caso del uso del aprendizaje automático. Se utilizará una pequeña cantidad de datos históricos para series de tiempo de ventas para el caso de que se proponga un nuevo producto o tienda. En la actualidad, se han desarrollado diferentes modelos de series temporales, por ejemplo, por Holt-Winters,

ARIMA, SARIMA, SARIMAX, GARCH, etc. algoritmos para series de tiempo. en este estudio se consideró como datos históricos el período de tiempo largo (2 años) y para un período de tiempo corto (3 días) una tienda específica. Para nuestro caso de estudio, consideramos diferentes enfoques de aprendizaje automático para pronosticar series de tiempo. El efecto de la generalización del aprendizaje automático consiste en el hecho de capturar los patrones en la totalidad del conjunto de informaciones. Este efecto se puede emplear con el propósito de hacer predicción de ventas cuando hay una pequeña cantidad de datos históricos para ello ha realizado la construcción de conjuntos de regresión de modelos individuales tales como xtraTree, ARIMA, Random Forest, Lasso, Neural Network En nuestro caso de estudio, consideramos diferentes enfoques de aprendizaje automático para el pronóstico de series de tiempo. Considerando que la predicción de ventas es más un problema de regresión que un problema de series de tiempo. El uso de la regresión para la previsión de ventas a menudo puede brindarnos mejores resultados en comparación con los métodos de series de tiempo. Para ello se utilizará un enfoque de apilamiento de modelos de aprendizaje automático que al tener diferentes modelos predictivos con diferentes conjuntos de características, es útil combinar todos estos en resultados en uno. en este caso los resultados obtenidos Extra Tree, Arima, Random Forest, Lasso, Neural Network el error de validación (14.6%, 13.8%, 13.6%, 13.4%, 13.6%) y error fuera de muestra (13.9%,11.4%,11.9%,11.5%,11.3%) el resultado de apilamiento en el error de validación (12.6%) y el error fuera de muestra (10.2%). Los resultados muestran que el uso de apilamiento técnicas, podemos mejorar los modelos predictivos en cuanto al rendimiento para el pronóstico de series de tiempo de ventas.

Giri et al. (2019), realizaron la investigación, Forecasting New Apparel Sales Using Deep Learning and Nonlinear Neural Network Regression, en Japón. Pronóstico de las ventas de ropa nueva mediante el aprendizaje profundo y la regresión de redes neuronales no lineales. El comercio minorista de moda enfrenta muchos desafíos para prever la demanda futura de sus productos. Esto se debe a las elecciones siempre cambiantes de sus consumidores, para

predecir las preferencias futuras de los consumidores con anticipación; carecen de herramientas avanzadas para lograr este objetivo. Para superar este problema, este trabajo de investigación combina información histórica de productos con sus características de imagen utilizando el aprendizaje profundo y predice las ventas futuras. El objetivo de este trabajo es predecir la cantidad de nuevo producto utilizando datos históricos de ventas de ropa femenina ropa y sus correspondientes imágenes. Las Imágenes de ropa se convierten en vectores de características y luego se fusionan con datos históricos de ventas. Aplicamos la propagación hacia atrás neuronal modelo de red para predecir las ventas de un nuevo producto. Para lograr este objetivo, utilizamos un modelo de inicio de red neuronal profunda V3 para extraer los vectores de características de las imágenes. Cada imagen define un elemento único y se llevó a cabo en imágenes de ropa y sus ventas. Los datos de información fueron de 320 imágenes de ropa de mujer. Antes de fusionar los datos se dividió en entrenamiento y prueba considerando el 90% de los ítems para la formación y el 10% de los elementos para la prueba, es decir, 289 imágenes para entrenamiento y 31 imágenes para prueba. Después de fusionar la función de vectores de imágenes de productos con datos de ventas, el número final de observaciones fue 7403 y 748 en datos de entrenamiento y prueba, Se aplicó el "modelo MLP de red neuronal" en los datos para tres valores " α " para obtener el rendimiento óptimo del modelo y también para ilustrar el problema de sobreajuste. La predicción de los datos de entrenamiento es positiva (aproximadamente el 40% de variación) mientras que para los datos de prueba muestra la inversa relación, que es negativa. MSE y RMSE son dos indicadores populares en estadísticas para acceder al rendimiento de los modelos de regresión el rendimiento de la predicción en los datos de prueba para $\alpha = 0,0001$ que no está muy lejos del valor de los datos el valor RMSE es aceptable datos. En general, el mejor resultado de predicción es el valor alfa= 0,0001 para la agrupación de datos entrenado y de prueba dado el modelo rendimiento en datos entrenados Se puede interpretar fácilmente; más cerca de cero es el mejor. Entonces, podemos ver que el MAE para los datos de prueba es 13.197 y podría mejorarse aumentando el tamaño de datos de entrenamiento. El

enfoque novedoso para pronosticar las ventas de productos de moda basado en el aprendizaje profundo y la regresión NN no lineal es que se puede estimar el modelo con relaciones arbitrarias entre las variables independientes y las dependientes presentado en este documento. Los resultados de este estudio parecen prometedores para pronosticar la cantidad de ventas futuras de los productos.

Nunnari & Nunnari (2017), realizaron la investigación, *Forecasting Monthly Sales Retail Time Series: A Case Study*, en Grecia, Pronóstico de las ventas mensuales de la serie temporal minorista: un estudio de caso. En este informe los autores presentan un caso de estudio acerca del pronóstico de las series de tiempo mensuales del comercio minorista de ventas de varios artículos registrados por el censo de EE. UU. Bureau de 1992 a 2016. El problema de la modelización se aborda en dos pasos. Primero, las series de tiempo originales se eliminan de la tendencia mediante el uso de un enfoque de promediado. Posteriormente, el residuo las series de tiempo están modeladas por Auto-Regresivo No Lineal (NAR) cada serie temporal consta de alrededor de 300 muestras mensuales de tiempo incluyen bienes esenciales y no esenciales. para ello se utilizado dos tipos de enfoques El enfoque Neuro Fuzzy, donde uno de las particularidades más llamativas es que una vez que la red neuronal ha sido entrenada mediante el uso de algoritmos de aprendizaje automático, el modelo obtenido se puede interpretar en términos de una base de reglas y el enfoque de red neuronal FeedForward que consisten en una serie de neuronas artificiales simples, organizadas en al menos tres capas. La primera capa tiene una conexión desde la entrada de red y la capa tiene una conexión con la capa anterior siendo la capa final produce la salida de la red. Para evaluar objetivamente el desempeño de los modelos de pronóstico, se realizó el cálculo de varios índices, considerando tres índices de error, a saber, el sesgo, el MAE (error absoluto medio) y RMSE (error medio de la raíz cuadrada), se consideró las series de tiempo con un periodo de 12 meses para el propósito de la identificación del modelo, para cada serie temporal, los datos se dividieron en dos conjuntos adyacentes, denominados el entrenamiento y el conjunto de

pruebas, respectivamente el pronóstico se asumió en el rango $h \in [1, 12]$ meses. La identificación del modelo se realizó después de escalar la serie de tiempo residual en el rango $[-1, 1]$ para permitir una comparación directa entre los índices de rendimiento obtenidos para diferentes series de tiempo. Los resultados obtenidos fueron $mae \leq 0.22$, $RMSE \leq 0.36$ y $habilidad \geq 0.2$ para todas las series de tiempo consideradas y para un amplio rango del horizonte de pronóstico ($h \leq 10$). Un análisis más detallado muestra que entre las 12 series temporales minoristas consideradas las mejores los modelados son artículos deportivos, electrónicos y electrodomésticos, ropa, cerveza, computadoras, comida y autos nuevos, que exhibe para $h = 1$ $mae \leq 0,15$, $RMSE \leq 0,22$ y $habilidad \geq 0,5$. En conclusión los resultados del estudio de sobre el pronóstico y modelado de series de tiempo de ventas minoristas utilizando dos enfoques basados en redes neuronales. Evaluando los términos de MAE, RMSE, sesgo y habilidad, muestran que, a pesar de la escasa autocorrelación de la serie temporal residual a corto plazo, es beneficioso utilizar este tipo de modelos.

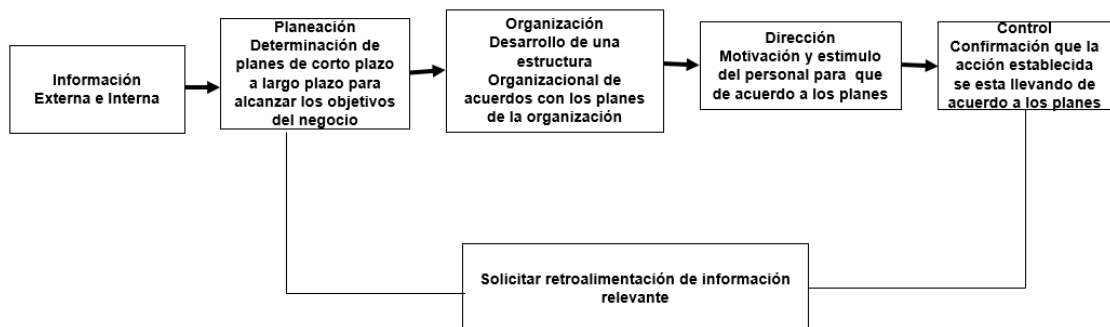
1.3. Teorías relacionadas al tema

1.3.1. El modelo de la gestión y la información

Actualmente en estos tiempos de cambios y sobre todo en de la tecnología, la relevancia de la información de forma procesada para la gerencia y el administrador es muy importante y trascendental, tanto así que ha llegado a cristalizarse en un recurso para las organizaciones y a nivel empresarial porque es incluido para los logros y metas de las organizaciones, lo que es sinónimo de: “el procedimiento de repartir los ingresos de una compañía, verbigracia, los recursos económicos y humanos se realizan por intermedio del control, organización, dirección y planeación, con la finalidad de elaborar las salidas, verbigracia, servicios y bienes requeridos por la clientela, tal que se logren las metas y propósitos de la compañía. Conllevando dicho proceso en un ambiente en constante evolución (Espinosa, 2020).

Figura 8

Información y las funciones gerenciales



Nota. Fuente: Espinoza (2020).

1.3.2. La Importancia de la Información

La información que forma parte de una salida lógica de cualquier sistema forma parte vital y de gran relevancia para las gerencias de cualquier tipo de organización con la finalidad de lograr la consecución de los objetivos organizacionales, ya sean estos a largo, medio o corto plazo. Las gerencias de negocios siempre optan por realizar una miscelánea de ingresos por unidad, costes, rotación, costos y precios de ventas, lo que, a las postrimerías del proceso obtienen la recuperabilidad más alta por sobre los capitales invertidos inicialmente. Esto es obtenido de informaciones adecuadas acerca de los acontecimientos esenciales, las gerencias pueden ampararse con mayor ahínco en métodos deductivos y analíticos que en los pensamientos intuitivos y las estimaciones, las mismas a las que normalmente suele estar forzada a hacer usanza, sobre todo cuando no hay existencia de informaciones de carácter relevante. Una gran cantidad de decisiones equívocas fueron el producto de informaciones inadecuadas, insuficientes o mal procesadas (Espinoza, 2020).

1.3.3. Inteligencia de Negocios (BI)

BI es aquella forma de descubrir informaciones oportunas de los datos de tal forma que se puede expresar y averiguar qué ha sucedido y porque ocurrió, la

BI consiste en una agrupación de herramientas que se utilizan con el propósito de recuperar y analizar informaciones de una determinada compañía o institución sobre el funcionamiento de la misma donde ayuda a lograr la consecución de conocimientos amplios de factores que afecta en su desenvolvimiento (logística, operaciones, producción, ventas, etcétera) y de este modo, seleccionar las decisiones correctas y lograr infundir tener los resultados más adecuados y mejores a la vez. En la praxis, implantar esta conceptualización tan moderna como lo es el BI involucra poseer una visión holística de la totalidad de las informaciones de la compañía que necesita de ella. Asimismo, se fundamenta en la usanza de estas informaciones con el propósito de promover innovaciones, eliminando y evitando ineficiencias y adaptándose raudamente a los cambiantes requerimientos y mercados actuales.

Es menester considerar que actualmente las soluciones modernas de BI están basadas al servicio de TI. El recojo y procedimiento de procesar informaciones producen datos y es por ello por lo que el cerebro humano se logra cristalizar hasta convertirse en inteligencia considerando un ámbito específico para un caso particular o, en su defecto, para cada individuo. “El procedimiento que logra el nacimiento de inteligencia propiamente dicha es el recogimiento continuo con confirmación, revisión y análisis de dichas informaciones permitiendo la comprensión de la problemática o el estatus de un modo actuable en concordancia con un usuario tipo final” (Gonzalez, 2015, pág. 19).

En el mundo actual que es altamente competitivo, la precisión y calidad de las informaciones de los giros de los negocios no es tan solo una opción o elección entre pérdida o utilidad, por el contrario, más bien, es catalogada como cuestión de conservación de la organización en el ámbito en el que se desenvuelve. Los analistas BI en cualquier industria señalan que, durante los venideros años, personas por millones harán usanza de BI en diversas técnicas y herramientas de análisis y visualización en los quehaceres diarios.

Las organizaciones que lideran sus diversos ámbitos alrededor del mundo globalizado en el que nos encontramos se encuentran desarrollando algo, en

detrimento de solamente adquirirlo y almacenarlo como, verbigracia, las informaciones con las que cuentan, por ello es que, se encuentran en construcción de estrategias competitivas, las cuales se encuentran fundamentadas en los conocimientos adquiridos por parte de las informaciones, lo que, valgan verdades, les viene entregando resultados efectivos y sobresalientes, y es utilizado como un as bajo la manga. Pues bien, para eso es el BI: modelamientos predictivos fundamentados en informaciones verídicas, tecnologías de información y sofisticados análisis estadísticos de naturaleza cuantitativa (Davenport & Harris, 2007). Todo esto se relaciona con el propósito de originar una vista holística de cualquier compañía y coadyuvar a los individuos a una toma decisional más útil, adecuada y mejor.

BI puede ser conceptualizado a manera de aquellas tecnologías y a aplicaciones de las que se hacen usanza con el propósito de obtener, lograr acceder y analizar informaciones y/o datos acerca de las operaciones desplegadas en una determinada compañía, las mismas que coadyuvan a la obtención de amplios conocimientos de aquellos elementos que terminan por afectar y, por consiguiente, aminorar, sus desempeños y, por ende, con esa obtención de informaciones poder lograr una óptima toma de decisiones para la mejora de dicha compañía en todos los ámbitos en los que se desenvuelve. Davenport & Harris (2007) revelaron que, BI es obtener, gestionar, tabular y reportar informaciones en pro de mejorar la toma de decisiones, considerando procedimientos computacionales y técnicas estadísticas-analíticas de las cuales se echa mano para el aprovechamiento y análisis de las mismas.

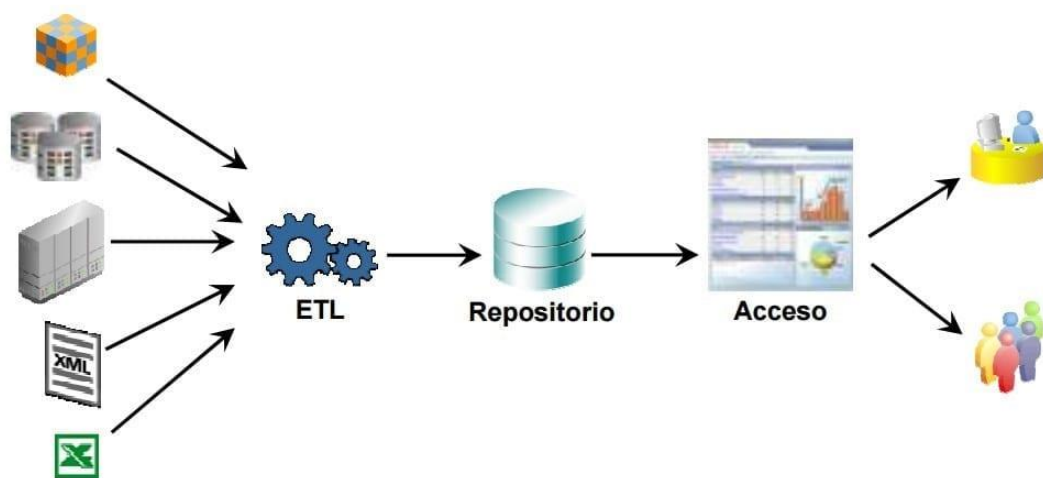
Por otro lado, Jourdan et al. (2008) Revelaron que, BI es tanto productos como también proceso, ambos a la vez. El proceso puesto que, se encuentra constituido por modelos, metodologías u otros de los cuales las compañías hacen usanza con el propósito de desplegar informaciones aplicables de manera inteligente, las cuales coadyuvan a dichas compañías en seguir el camino hacia la salida en un mundo que cada vez más se pone muy globalizado y competitivo. Por otra parte, desde el punto de vista de producto, es data que coadyuva a las compañías a lograr la predicción de comportamientos emanados por parte de la competencia, acerca de clientes, de proveedores, de TI, del

proceso adquisitivo, de los movimientos de los mercados, de la prestación de servicios o productos y de la conducta en general de los diferentes ámbitos de los negocios, considerando para esto un grado de exactitud y puntualidad en dichas informaciones.

El concepto de BI se puede rastrear desde los años 70, con aquellos reportes desarrollados originariamente por parte de MIS (por su acrónimo en inglés de Management Information Systems), los mismo que, por aquel entonces mostraban mucho estatismo y simplicidad pero que, posteriormente, en la década siguiente permitió el surgimiento del concepto de EIS (por su acrónimo en inglés de Executive Information Systems), el mismo que involucraba particularidades multidimensionales, dinámicas y características propias de la coyuntura de tiempo actuales, con algunos elementos que sirvan para la realización de reportes, pronósticos, analíticas de informaciones, etcétera. Posteriormente, en la década de los 90, el Grupo Gartner INC ya acuñó la conceptualización de BI. Finalmente, BI, desde el 2005 a la fecha circunscribe asimismo a la denominada Inteligencia Artificial, así como también a otros elementos analíticos con gran potencial. Las técnicas y herramientas que forman parte indispensable de BI pueden visualizarse en la figura siguiente:

Figura 9

Técnicas y herramientas de BI



Nota. Fuente: Turban et al. (2008).

Minería de datos: es el proceso de utilizar BD, estadísticas y correlaciones en un conjunto de datos.

Generación de informes: consiste en mostrar y distribuir la analítica de informaciones con los stakeholders de modo que, éstos a su vez puedan conseguir obtener sus conclusiones propias, para su toma decisional.

Valores de referencia y métricas de rendimiento: La comparación entre los datos históricos de la compañía con las informaciones del actual rendimiento permiten la realización de seguimientos de los rendimientos frente a las metas organizacionales.

Análisis descriptivos: Hacer usanza de analítica a las informaciones preliminares con el propósito de saber a ciencia cierta qué ocurrencias o acontecimientos se vienen dando.

Consultas: los usuarios pueden plantear pesquisas específicas las cuales deben de estar interrelacionadas con BI y la data recapitulando las respuestas de las agrupaciones de informaciones.

Análisis estadístico: a raíz de las informaciones vertidas gracias al análisis descriptivo, dichas informaciones son exploradas con mayor profundidad mediante la usanza de estadísticas, verbigracia, para lograr establecer una tendencia de ventas, como es que ocurre y porqué es que viene sucediendo.

Visualización de datos: es la transformación de los análisis hechos a las datas informativas con el propósito de cristalizarlos en reportes, o en cualquier otra representación, ya sean mediante histogramas, gráficos u otro visual.

Análisis visual: explora las informaciones por intermedio de la narrativa visual de hechos con el propósito de compartir informaciones acerca de cómo se debe de seguir, permaneciendo siempre en un análisis fluido de dichas informaciones.

Preparación de datos: recapitular diversas fuentes informativas, con el propósito de examinar y reconocer las medidas y las dimensiones para, a

posteriori, realizar la preparación de información para análisis estadístico cuantitativo.

1.3.4. Sistema de Soporte de Decisiones e Inteligencia de Negocios

Un sistema de soporte de decisiones (DSS) es un sistema de información interactivo que analiza grandes volúmenes de datos para informar decisiones comerciales (Gorry & Morton, 1971). Un DSS respalda los niveles de administración, operaciones y planificación de una organización para tomar mejores decisiones al evaluar la importancia de las incertidumbres y las compensaciones involucradas en la toma de una decisión sobre otra.

Un DSS aprovecha una combinación de datos sin procesar, documentos, conocimiento personal y/o modelos comerciales para ayudar a los usuarios a tomar decisiones. Las fuentes de datos utilizadas por un DSS podrían incluir fuentes de datos relacionales, cubos, almacenes de datos, registros médicos electrónicos (EHR), proyecciones de ingresos, proyecciones de ventas y más (Turban, Sharda, Aronsson, & King, 2008).

El concepto de sistemas de apoyo a las decisiones surgió a partir de una investigación realizada en el Carnegie Institute of Technology en las décadas de 1950 y 1960, pero realmente echó raíces en la empresa en la década de 1980 en forma de sistemas de información ejecutiva (EIS), sistemas de apoyo de decisiones grupales (GDSS) y sistemas de apoyo a la toma de decisiones organizativas (ODSS) (Porter, 1980). Estos dicen, a medida que las organizaciones se centran cada vez más en la toma de decisiones basada en datos, la ciencia de decisiones (o inteligencia de decisiones) va en aumento, y los científicos de decisiones pueden ser la clave para desbloquear el potencial de los sistemas de ciencia de decisiones. Al reunir la ciencia de datos aplicada, las ciencias sociales y la ciencia de la gestión, la ciencia del diseño se centra en seleccionar entre opciones para reducir el esfuerzo requerido para tomar decisiones de mayor calidad.

Los sistemas de soporte de decisiones y la inteligencia empresarial (BI) a menudo se combinan. Algunos expertos consideran a BI un sucesor de DSS. Los sistemas de soporte de decisiones generalmente se reconocen como un elemento de los sistemas de inteligencia empresarial, junto con el almacenamiento de datos y la minería de datos.

Mientras que BI es una categoría amplia de aplicaciones, servicios y tecnologías para recopilar, almacenar, analizar y acceder a datos para la toma de decisiones, las aplicaciones DSS tienden a estar más diseñadas para respaldar decisiones específicas. Por ejemplo, un DSS comercial puede ayudar a una empresa a proyectar sus ingresos durante un período determinado mediante el análisis de datos de ventas de productos anteriores y variables actuales. Los proveedores de atención médica utilizan sistemas de apoyo a la toma de decisiones clínicas para hacer que el flujo de trabajo clínico sea más eficiente: alertas y recordatorios computarizados para los proveedores de atención, pautas clínicas, conjuntos de órdenes de afecciones específicas, etc.

Según Imhoff & Pettit (2004) los sistemas de apoyo a la toma de decisiones se utilizan en una amplia gama de industrias. Los usos de ejemplo incluyen:

- a. Planificación de rutas GPS. Se puede utilizar un DSS para planificar las mejores y más rápidas rutas entre dos puntos analizando las opciones disponibles. Estos sistemas a menudo incluyen la capacidad de monitorear el tráfico en tiempo real para evitar la congestión.
- b. Planificación de cultivos. Los agricultores usan DSS para ayudarlos a determinar el mejor momento para plantar, fertilizar y cosechar sus cosechas. Bayer Crop Science ha aplicado análisis y apoyo a la toma de decisiones en todos los elementos de su negocio, incluida la creación de "fábricas virtuales" para realizar análisis hipotéticos en sus plantas de fabricación de maíz.
- c. DSS clínico. Estos sistemas ayudan a los médicos a diagnosticar a sus pacientes. Penn Medicine ha creado un DSS clínico que ayuda a que los pacientes de la UCI dejen de usar ventiladores más rápido.

- d. Cuadros de mando ERP. Estos sistemas ayudan a los gerentes a monitorear los indicadores de desempeño. La empresa de servicios y marketing digital Clearlink utiliza un sistema DSS para ayudar a sus gerentes a identificar qué agentes necesitan ayuda adicional.

Según Olavsrud (2020) Entre los principales softwares del sistema de apoyo a la toma de decisiones tenemos los siguientes:

- a. Information Builders WebFOCUS. Esta plataforma de datos y análisis está diseñada para empresas y empresas medianas que necesitan integrar e incrustar datos en todas las aplicaciones. Ofrece opciones en la nube, múltiples nubes, locales e híbridas (Olavsrud, 2020).
- b. QlikView. QlikView es la solución analítica clásica de Qlik, construida sobre el motor asociativo de la empresa. Está diseñado para ayudar a los usuarios con sus tareas diarias mediante un panel configurable (Olavsrud, 2020).
- c. SAP BusinessObjects. BusinessObjects consta de aplicaciones de análisis e informes para ayudar a los usuarios a comprender las tendencias y las causas fundamentales (Olavsrud, 2020).
- d. TIBCO Spotfire. Este software de visualización y análisis de datos ayuda a los usuarios a crear paneles y potenciar aplicaciones predictivas y aplicaciones de análisis en tiempo real (Olavsrud, 2020).
- e. Salesforce Analytics Cloud. Esta solución de análisis basada en la nube impulsada por inteligencia artificial se basa en la plataforma de Salesforce.com para ayudar a las organizaciones a detectar oportunidades y predecir resultados (Olavsrud, 2020).
- f. Powernoodle. Powernoodle es una plataforma de participación de decisiones basada en la nube que aprovecha la ciencia cognitiva, conductual y de decisiones. Ofrece plantillas prediseñadas que abordan tipos de decisiones comunes y soporte para modelar los flujos de trabajo de múltiples grupos de partes interesadas (Olavsrud, 2020).

- g. 1000minds Decision Making. 1000minds es un conjunto de herramientas y procesos en línea para la toma de decisiones, la priorización y el análisis conjunto. Se deriva de la investigación realizada en la Universidad de Otago en la década de 1990 sobre métodos para priorizar pacientes para la cirugía (Olavsrud, 2020).
- h. Briq. Briq es una plataforma de automatización y análisis predictivo construida específicamente para contratistas generales y subcontratistas en la construcción. Aprovecha los datos de contabilidad, gestión de proyectos, CRM y otros sistemas para potenciar la inteligencia artificial para el análisis predictivo y prescriptivo (Olavsrud, 2020).

1.3.5. Implementación de BI

En un sistema de BI se debe tener en cuenta tipos posibles de usuarios potenciales, para la alineación de la estrategia de negocios. Donde le permite a las compañías desplegar sus operaciones, logrando el mejoramiento de sus procesos, y logrando una toma decisonal en concordancia con los datos e informaciones obtenidas. Asimismo, es menester disponer de una estructura o esquema de formación de la proyección y la puesta en marcha de las funciones siguientes: negocios, organizaciones, funcionalidades, e infraestructuras. Por lo permite definir los objetivos estratégicos y operacionales considerando, asimismo, las habilidades de la compañía y la cultura de la empresa, y esto permite interesar a los grupos de trabajo para lo lograr los objetivos. Si la totalidad de ellos están en alineamiento respectivo, al poner en marcha la implantación del sistema BI, es menester implementar un Centro de Competencia BI, el mismo que se convierta en soporte para la totalidad del proceso (Turban, Sharda, Aronsson, & King, 2008).

En la actualidad, BI se ha convertido en un socio importante para las estrategias de negocios, para lograr el desempeño de manera competitiva, y actualmente hacen usanza de estas las compañías de cualquier tamaño, sin importar la cantidad de colaboradores con las que cuentan ni su capital. Algunos autores recomiendan que BI debería convertirse en un término difundido en completitud

en todas las compañías, de esta manera la totalidad de trabajadores lograrían aprovechar en gran modo y se logre la consecución de más y mejores resultados con la prioridad de bajar los costos. Para realizarlo de manera óptima en una empresa cuando se implementa este debe ser integrado en toda la empresa, permitiéndole el mejoramiento del rendimiento de la compañía (Baum, 2006; Lal, 2005).

Angelo (2006) nos da a conocer que la Inteligencia de Negocios (BI) es proveída por muchos proveedores a manera de aplicaciones informáticas las cuales han sido construidas, verbigracia, CRM o ERP, las cuales en varios casos cuentan con soluciones internas de BI muy bien especificadas y desarrolladas, empero, también existen otros software como los antes mencionados que no cuentan con dichas particularidades de BI, y no dan a conocer los resultados esperados por las empresas. BI logra sorprendentes resultados para un eficiente manejo de las compañías. Una apuesta a posteriori, conveniente talvez, sería que estos proveedores de soluciones BI coadyuven en labores que permitan la difusión de las ventajas y provechos de BI y ofrecer diversos casos de éxito mostrando resultados palpables y fidedignos acerca de dichas implementaciones exitosas.

1.3.6. ¿Por qué es importante BI?

BI permite a las organizaciones un mejoramiento en la toma de decisiones, porque ofrece datos históricos y actuales en el ámbito empresarial en el que se desenvuelven, siendo de mucha ayuda para los diversos analistas de un abanico de áreas de modo que puedan valerse de BI con el propósito de establecer valores referenciales de la competencia y del rendimiento que coadyuven a dichas compañías a desarrollar sus actividades de modos más eficaces, eficientes y ágiles. Como también, mediante esta BI podrán lograr la identificación de marcas actuales en cuanto al mercado y sus tendencias de un modo más fácil con el propósito de lograr el mejoramiento e incremento de sus niveles de ingresos y/o ventas, mejorando de esta manera su rentabilidad. Permitted mejorar todo tipo de iniciativas para, inteligentemente y basado en

dichas informaciones, desarrollar una mejor toma de decisiones (Grupo Salesforce, 2021).

- Reconocer modos para lograr el aumento de beneficios
- Analizar los comportamientos de la clientela
- Predecir el éxito
- Realizar una comparación de las informaciones con los datos de los competidores
- Detectar problemáticos o inconvenientes
- Verificar un monitoreo y seguimiento de la performance
- Optimizar las operaciones
- Señalar las conductas y tendencias de los mercados

1.3.7. Data Analytics

Es una evolución de etapas o tareas que pretende identificar la necesidad de un negocio o empresa buscando como solución esta necesidad contando con el apoyo primordial de individuos que cuentan con habilidades para analizar datos apoyándose de técnicas y/o herramientas que tengan criterios de para establecer métricas las cuales se consideran para evaluar si son válidas para dicha solución. La tendencia actual para laborar en posiciones que se encuentren enfocadas en el análisis de datos se encuentra en aumento puesto que, las compañías que laboran de este modo toman decisiones operativas más óptimas y gestionan mucho mejor sus riesgos obteniendo un crecimiento rentable y sostenible. Actualmente un gran número de compañías efectúan inversiones en aplicaciones TI, empero carecen del capital humano idóneo dado que ellos están faltos de entrenamientos y capacitaciones apropiadas. Haciendo la pregunta: el Data Analytics ¿Para qué sirve? en concreto es aquel que impacta en toda la cadena de valor de las organizaciones (Digital House, 2021). A manera de ejemplificación, es útil para:

- Identifica y anticipar riesgos
- Optima las fuerzas de venta
- Estima previsiones en cuanto a la producción
- Identificar los clientes y productos con menores niveles de rentabilidad.

Análisis Predictivo

El análisis predictivo consiste en emplear métodos estadísticos a través de algoritmos, métodos y técnicas de forma automática con la finalidad de identificar comportamientos y/o probabilidades de resultados futuros a partir de información y datos históricos.

Este análisis reside en la extracción de un modelo analítico de informaciones históricas, las cuales predicen comportamientos o tendencias futuras, o en su defecto, estiman resultados poco conocidos haciendo usanza de herramientas y técnicas de carácter estadístico matemático para el modelamiento de modo que, reúna la totalidad de información permitiendo de esta manera la realización de análisis predictivos a futuro.

La minería de datos en conjunto con la estadística, admiten la construcción de modelos predictivos, descubriendo correlaciones y tendencias, por intermedio de agrupaciones de datos no estructurados como estructurados (Big Data Marketer, 2018).

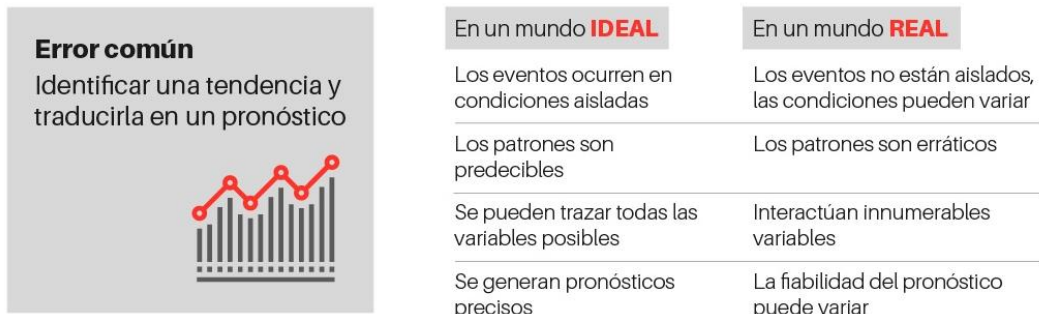
Analítica Predictiva

Cuando se aplica operaciones avanzadas de matemáticas, verbigracia, algoritmos automatizados o análisis estadístico, los cuales brinden patrones estadísticos logrando adelantarse al futuro, prediciendo resultados, las compañías se encuentran con un abanico de posibilidades característicos de la analítica predictiva, las cuales admiten adelantarse al futuro y conocer tendencias, conductas y comportamientos de diversos indicadores claves y valiosos para el giro del negocio de distintas compañías. Y la vez nos brindan

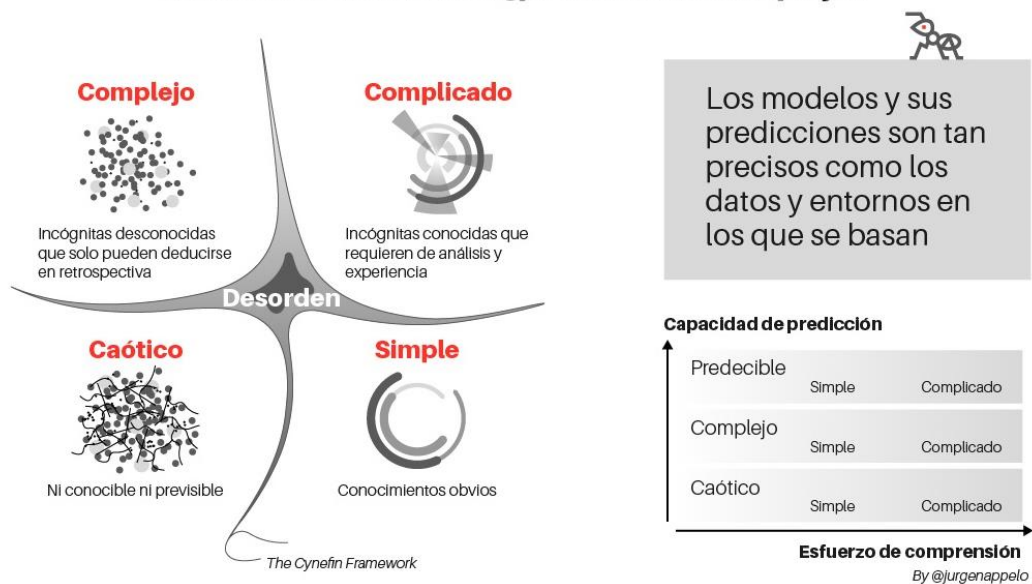
información valiosa por intermedio de los datos con el propósito de coadyuvar a las compañías e instituciones a tomar decisiones adecuadas para los negocios (Decide, 2021).

Figura 10

Proceso de analítica predictiva



Relación causa-efecto: ¿predecible o compleja?



En sistemas complejos la probabilidad es limitada



Nota. Fuente: Muñoz (2019).

Como se sabe, actualmente a nivel mundial en las empresas manufactureras e industrias realizan la producción de cualquier bien de consumo donde supervisan los niveles de servicios internos/externos y los inventarios. Estas compañías manufactureras deben de tomar, a diario, decisiones acerca de tiempos de entregas, de volúmenes de producción, de entregas de embarques, de tiempos muertos por cada uno de sus procesos, considerando la disponibilidad de sus recursos, verbigracia. A lo largo de la planificación, deben de tenerse en consideración diversidad de factores tales como la utilización de recursos humanos e industriales, las coberturas de stocks, permitiendo la eficacia en las campañas de las líneas de producción y de distribución de los productos (Decide, 2021)

Se pone como ejemplo la campaña presidencial desplegada por Barack Obama en su candidatura, en el 2008, a la presidencia de los Estados Unidos, en la cual se hizo de modeladores predictivos y personal estadístico con experiencia en extracción de datos, además de analistas, programadores y matemáticos, con el propósito de obtener datos e informaciones de votantes potenciales en cada uno de los estados que componen dicha nación y, de este modo, evaluar que metodología podría ser la ideal para lograr un acercamiento hacia estos potenciales votantes con el propósito de ganarse su voto mediante la exposición de sus propuestas de gobierno, logrando un éxito abismal lo que le dio la oportunidad de volver a emplear dichas tácticas para su campaña presidencial del año 2012, en la cual, en dicha campaña se logró una cifra de 5 veces en comparación con la campaña antes mencionada.

El término de analítica se puede definir de la manera siguiente: Procedimiento de carácter científico que tiene como propósito transformar los datos en informaciones, y estas a su vez en conocimientos que permitan tomar decisiones más acertadas e idóneas (Higuera, 2016). Asimismo, por añadidura, se sabe de la existencia de múltiples tipologías de analíticas, empero, las tres (03) más conocidas se conceptualizan en el siguiente diagrama:

Figura 11

Clasificación de la analítica



Nota. Fuente: Higuita (2016).

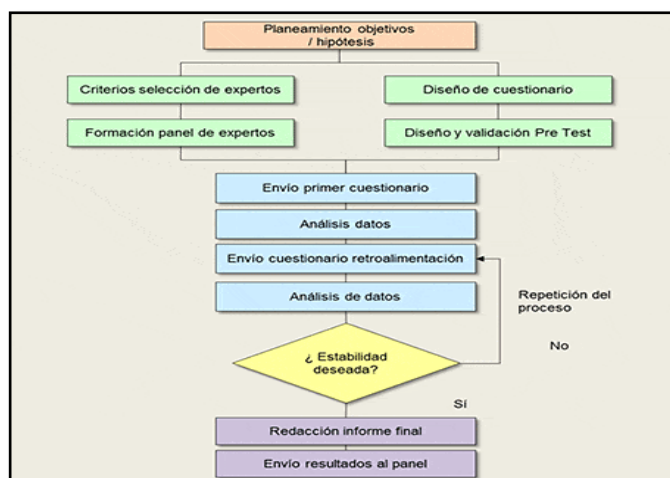
1.3.8. Modelos de Predicción

Método Delphi

Este método pertenece forma parte de la técnica cualitativa de pronóstico de ventas es uno de lo que más se utiliza consiste en varias etapas en la cual se considera un grupo de expertos, la cual se encuentra conceptualizada en la Figura 5:

Figura 12

Etapas del Método Delphi



Nota. Fuente: Cruz & Rúa (2018).

- **Etapa 1.** Cada persona que integra el grupo brinda una respuesta escrita sobre las preguntas realizadas en el cuestionario propuesto. Para luego de estas respuestas se tabulan y se distribuyen a todo el grupo de expertos para poder calificar dichas respuestas que se basan en valores estadísticos como promedios valores máximos y mínimos desviación estándar, etc.
- **Etapa 2.** En esta etapa las respuestas se vuelven a resumir y se da un feedback al grupo de expertos determinando a una conclusión.
- **Etapa 3.** se vuelven a resumir nuevamente las respuestas y vuelve a dar otro feedback al grupo de expertos y así sucesivamente se repite este proceso hasta alcanzar un grado suficiente de acuerdo de todos y estos utilizándolos como pronósticos en una última etapa. De esta forma las ideas y las percepciones de un grupo se consideran como naturaleza heterogénea donde se convergen de forma progresiva a consensos y conclusiones llegando a disminuir el riesgo asociado y la toma de decisiones en el pronóstico de ventas.

Análisis Bayesiano

También conocido como el Teorema de Bayes, es manipulado con el propósito de calcular probabilidades de diversos sucesos, considerando como ase información sobre ese suceso este teorema entiende las probabilidades de forma inversa al teorema de probabilidad total. Está conformado por dos variables.

- A : Que cumple cierta característica con condiciona su probabilidad
- B : Es la inferencia sobre un suceso en b a partir de los sucesos de A.
- N : Cantidad de Datos disponibles a Pronosticar

Por tanto, $P(A/B)$ son posibilidades a posteriori, $P(A)$ posibilidades a priori y $P(B/A)$ posibilidad de que se dé B en cada una de las hipótesis de A. Formula:

$$P[A_n/B] = P[B/A_n]P[A_n] / (\sum P[B/A_i]P[A_i])$$

Este teorema ha sido muy cuestionado debido a su mala aplicación para ello ha sido necesario que mientras los supuestos se cumplan sucesos exhaustivos y disjuntos, dicho teorema es válido en su totalidad.

Su utilidad ha demostrado que se basa en la experiencia concreta sobre algo las observaciones y evidencias que se hacen usanza para inferir y/o actualizar las probabilidades siendo de mucha de utilidad en diversos campos, desde la medicina, el juego de azar y hasta el tratamiento informático de datos (González J. , 2015).

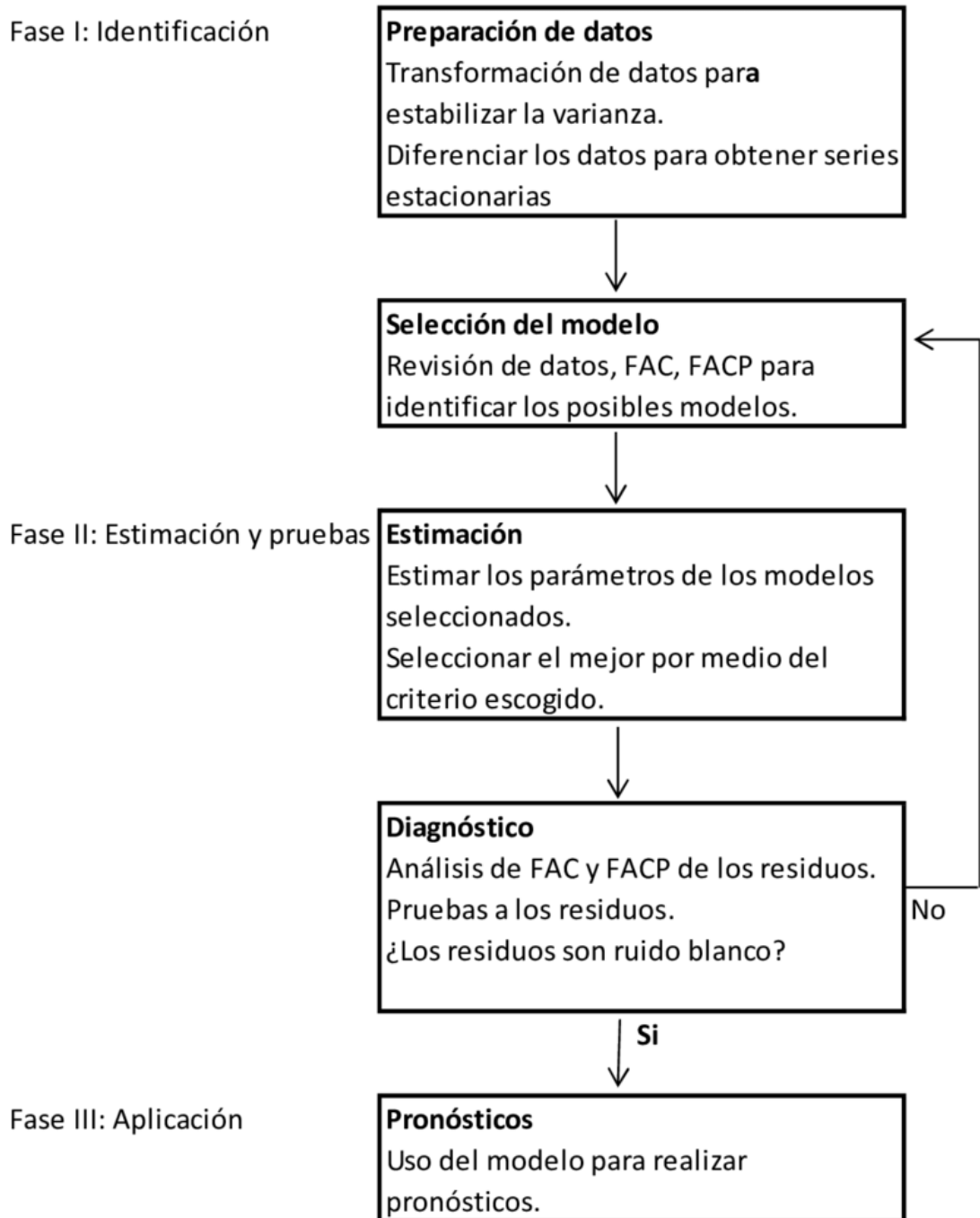
Modelo autorregresivo integrado de promedio móvil (ARIMA)

El modelo ARIMA, el cual fue por George Box y Gwilym Jenkins este modelo está compuesto de procesos acrónimo como de auto regresivos, integrados y media móviles es usado como una alternativa a la modelización y predicción tradicional mediante modelo estructurales consiste en aplicar series temporales que son generadas por un proceso de datos los cuales pueden ser identificados y cuantificados para ser inferidos sus valores al futuro.

Este proceso de datos que se componen como una serie de comportamientos, interactúa con procesos completamente aleatorios conocido como ruido blanco. Las etapas en la elaboración de un modelo ARIMA son: identificación, estimación, aplicación (validación y predicción).

Figura 13

Modelo de Aplicación ARIMA



Nota. Fuente: Díaz et al. (2018).

En modelo general ARIMA se compone de las variables (p,d,q) y es conceptualizado como un proceso auto regresivo integrado de medias móviles de orden p, d, q, tomando la expresión general de la ecuación siguiente:

$$(1 - a_1 L^1 - a_2 L^2 - \dots - a_p L^p) (1 - L)^d Y_t = (1 - b_1 L^1 - b_2 L^2 - \dots - b_q L^q) \varepsilon_t$$

Siendo:

Y_t : Valor de la variable en el momento t .

L : Operador de retardos ($Y_{t-1} = Y_t * L$)

a_p : Coeficientes del operador de retardos para el componente auto regresivo.

b_q : Coeficientes del operador de retardos para el componente de medias móviles.

ε_t : Componente aleatoria.

Asimismo, el término de error es denominado ruido blanco y tiene que ser de media nula, varianza constante y covarianza nula entre errores de observaciones diferentes esta serie temporal compuesta de (0, d, 0) después de ser diferenciada d veces.

Los modelos ARIMA (p, d, q) permiten describir las observaciones tras haber sido diferenciadas d veces para eliminar las posibles fuentes de no estabilidad. Si hay un componente igual a cero o es nulo y se excluye del término correspondiente de la fórmula general. Por ello es necesario que se cumpla la condición de estabilidad. En la serie temporal ha de ser al menos débilmente estacionaria. Si se quiere determinar el proceso aleatorio mediante un modelo de tipo auto regresivo es necesario asegurarse que la serie debe ser estacionaria, y si no lo es, hay que transformar la serie temporal original para obtener otra serie temporal que cumpla el criterio de invariabilidad con respecto al tiempo (Gallego-Nicasio, Rodríguez, Mínguez, & Jiménez, 2018).

Método Holt-Winters

Este método es una ampliación perfeccionada orientada a la suavización exponencial que brinda de forma general el análisis de la información a largo plazo que permite la elaboración de pronósticos a corto plazo bajo la conducta de una serie temporal que nace de las informaciones históricas obtenidas, siendo este método interactivo usado dado que permite analizar y pronosticar series de tiempos.

Este método tiene dos modelos dependiendo del tipo de estacionalidad son:

- a. **Modelo Multiplicativo estacional** que a medida que se incrementan los datos se incrementa el patrón estacional.
- b. **Modelo aditivo estacional** donde los factores se diferencian y se agrupan para modelar los datos.

Ambos modelos utilizan 3 ponderaciones o parámetros de suavización para actualizar los componentes en cada periodo. Los valores iniciales para los componentes de nivel y de tendencia se obtienen de una regresión lineal sobre el tiempo. Los valores iniciales para el componente estacional se obtienen de una regresión de variables simulada utilizando datos sin tendencia. En el caso de realizar pronósticos se utilizan datos del tipo de origen del pronóstico para generar pronósticos (Minitab, 2019).

Formula: **El pronóstico para m períodos adelante desde un punto en el tiempo t es:**

Método multiplicativo:

$$(L_t + mT_t) S_{t+m-p}$$

Método aditivo:

$$L_t + mT_t + S_t + m - p$$

Donde:

L_t : Nivel

T_t : La tendencia en el tiempo t

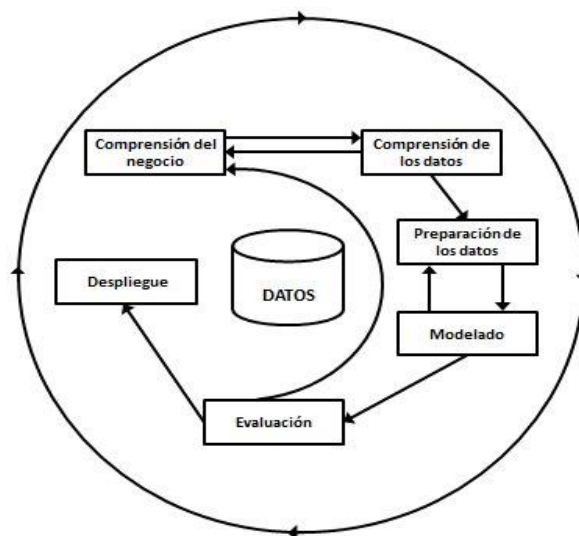
$S_t + m - p$: Componente estacional para el mismo período del año anterior)

Metodología CRISP – DM (Cross Industry Standard Process for Data Mining)

CRISP-DM Es la metodología que organiza el progreso del proyecto de una minería de datos en series de fases o etapas, con trabajos de forma general y específicos que permiten cumplir con los objetivos del proyecto. Dichas fases trabajan de forma cíclica e iterativa, pudiendo regresar de una fase a otra anterior. Se basa en función a un modelo jerárquico de procesos, donde se establece un ciclo de vida de los proyectos de explotación de información.

Figura 14

Fases del Modelo CRISP-DM



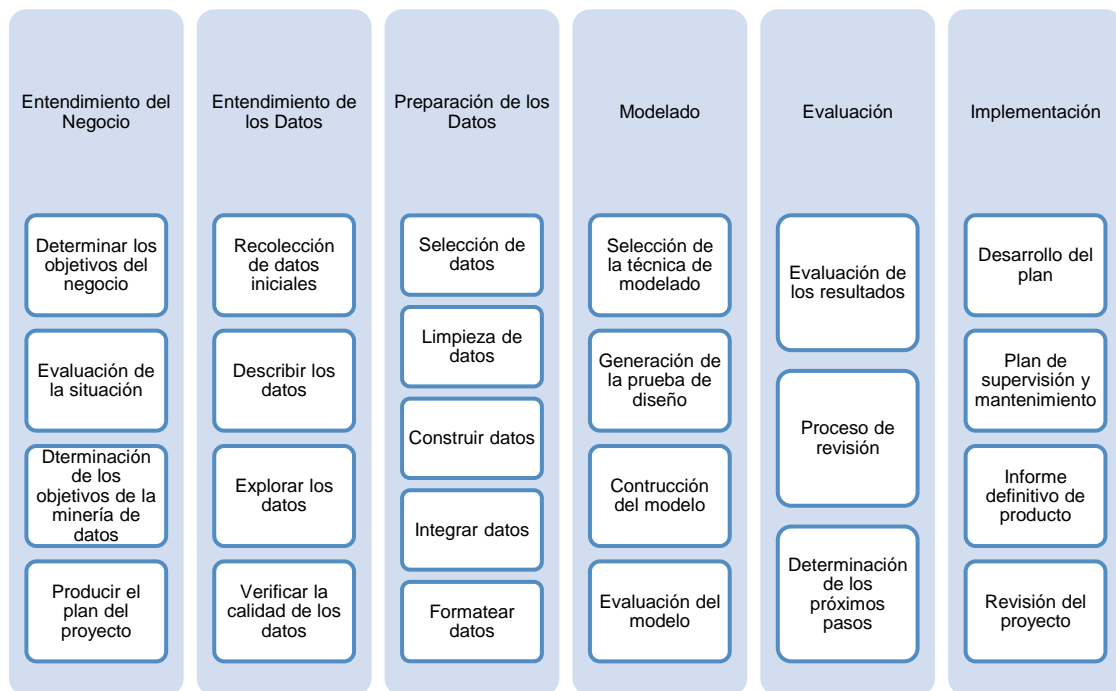
Nota. Fuente: Chapman et al. (2000).

Según Hernández-Orallo et al. (2021) las fases de la metodología CRISP son las siguientes:

- a. Comprensión del negocio:** Es donde trata de deducir los objetivos y requerimientos del proyecto desde una perspectiva de negocio.
- b. Comprensión de los datos:** En fase se eligen y adapta los datos, para poder identificar los problemas de calidad de datos y así obtener datos óptimos para poder ser analizados.
- c. Preparación de los datos:** Transformación de los datos. Se escogen los datos a trabajar y éstos pasan a una fase de limpieza, estructuración, integración y formateo.
- d. Modelamiento y evaluación:** Separación y aplicación de Data Mining e Interpretación y evaluación. Se selecciona la técnica a trabajar, construyendo el modelo, para luego ser sometido a diferentes pruebas y evaluaciones.
- e. Despliegue del proyecto:** Es donde se explota todo el potencial de los modelos y así intégralos en los procesos de toma de decisión de organización, y así difundir el conocimiento extraído, etcétera.

Figura 15

Tareas y subtareas del proceso CRISP-DM



Nota: Fuente: Chapman et al. (2000).

En general, las fases de CRISP-DM tienen una descripción sobre el proceso de análisis, pero los detalles específicos y las necesidades del usuario deben ser incluidas. Se necesita además de la actualización un mantenimiento constante para ser adaptado en los nuevos retos, como: big data, ciencia de datos, minería de datos centrada en el usuario.

1.3.9. El papel de los sistemas de información en las ventas

Los sistemas de información desempeñan un papel importante en la gestión de ventas. Son herramientas analíticas que permiten a la organización trazar ventas para el presente y futuro la información que brindan es de forma instantánea. La automatización de los procesos de ventas ayuda a reducir el trabajo en papel, ponen al alcance mejor información rápidamente y facilitan la tarea de enfocarse en los principales puntos. El software especializado está ayudando a perfeccionar la efectividad y la eficiencia en el mercado competitivo

hoy en día el desarrollo de nuevas estrategias se necesita estar interconectadas y contar con la información necesaria en el momento oportuno para facilitar la toma de decisiones. Los sistemas de información se deben adaptar las necesidades de cada empresa, sea ésta grande o pequeña, deben ser amigables y flexibles. Para definir algunos conceptos básicos que nos ayudarán a entender cómo funcionan los

Sistemas de Información (SI)

- a. Sistema**, Colección de objetos, procedimientos o técnicas que interactúan de manera regulada formando un todo organizado.
- b. Datos**, Las empresas cuentan con grandes cantidades de datos que por sí mismos no son información confiable; estos datos requieren un proceso para ser útiles y manejables.
- c. Información**, Carece de significado si es que se encuentra aislada y debe ser procesada mediante una clasificación, almacenamiento y relación para que luego pueda realmente informar sobre algún tema. La información es un recurso estratégico y un arma competitiva poderosa.
- d. Sistemas de Información**, Se definen como sistemas formales para recopilar, integrar, comparar, analizar y difundir información interna y externa de la empresa, en forma oportuna, eficaz y eficiente. Un sistema de información gerencial es un sistema útil para la toma de decisiones del personal gerencial de todos los niveles de la organización para llevar a cabo las funciones de gestión inherentes a su cargo.

1.3.10. Redes Neuronales

Las redes neuronales artificiales (RNA), es una técnica e la inteligencia artificial. Mediante la cual intenta desarrollar soluciones informáticas emulado el comportamiento del cerebro humano, actualmente estas son utilizadas en diferentes campos como neurociencias, matemáticas, estadísticas, ciencias de la computación, etc. Las redes neuronales Artificiales según (Callejas I. ,

Piñeros, Rocha, Hernandez, & Delgado, 2013). Son parte de la inteligencia artificial donde son entrenadas a través de entradas obtenidas a partir de escenarios externos o internos y estas se múltiple por pesos asignados al azar siendo capaces de resolver funciones no lineales en un corto tiempo porque aprende de los datos que son difíciles de expresar matemáticamente son "herramientas poderosas para el análisis de señales y la modelación de sistemas". Permitiendo resolver problemas de forma individual o combinada con otros métodos para aquellas tareas de clasificación, identificación y diagnóstico.

Tabla 1

Características de las RNA

Característica	Descripción
Topología de red	Permite determinar la capacidad representativa de como una cantidad de neuronas están distribuidos en capas y distribuidas entre sí. En donde la topología de red se basa en el problema a resolver.
Regla de aprendizaje	Una red neuronal es una integración de diversos sistemas de aprendizaje, es por ello que tienen la capacidad de aprender a través de un entrenamiento previo.
Tipo de entrenamiento	Una RNA presenta dos tipos de entrenamiento, una que durante la etapa de aprendizaje la red se entrena para que los pesos sinápticos se adecuen a la red. La segunda es una etapa de ejecución en donde la red ya pasa ser operativa cuya red toma un valor de funcionamiento real.

Nota. Fuente, Gestal (2014).

Ventajas

- Tiene un aprendizaje adaptable donde aprende a realizar tareas a partir de un conjunto de datos los cuales son representados como entradas y pesos.
- Ordenamiento creando su propia organización o representación de la información recibida durante el aprendizaje utilizando I métodos matemáticos Adeline, Madeline y Perceptrón entre otros.
- Operación a tiempo real que pueden ser aplicadas por computadoras o dispositivos de hardware especial aprovechando la capacidad de las RNA.

Desventajas

- Tienen complejidad para el aprendizaje en las tareas grandes
- Elevado Tiempo de aprendizaje dependiendo mucho del número de patrones a reconocer
- Tiempo de coincidencia es decir existen aplicaciones que requieren un enorme conjunto de ejemplos para el entrenamiento.

1.3.11. Formas de Aprendizaje

Aprendizaje supervisado

Es una técnica de las redes neuronales que se basan en deducir una función a partir de datos o información de entrenamiento, dichos datos están estructurados por vectores u objetos dados en par, y un par de la información es la entrada mientras que el otra son los resultados. Los pesos, que son establecidos de manera aleatoria en un principio, son ajustados por la red produciendo un resultado más cercano entre el valor esperado y la salida real. Hasta ajustar la salida con un margen de error aceptable. El objetivo de este sistema de aprendizaje es predecir de cualquier entrada valida que el supervisor ingrese (Tang, 2014).

Aprendizaje no supervisado

Es que aquel cuando el proceso de aprendizaje no es inspeccionado ni controlado por ningún agente externo. Durante el proceso de aprendizaje a la red no se la presenta la salida deseada. Sus principales utilidades son, entre otras, descubrir las regularidades presentes en los datos, extraer rasgos o agrupar patrones, según la similitud a las redes de entrada para esto necesita una gran cantidad de vectores de entrada de tal manera que dispone de números suficientes de patrones obteniendo resultados de calidad (Flórez, 2008)

1.3.12. Machine Learning

Es conocida como una disciplina de la inteligencia artificial que a través de algoritmos dota a los ordenadores con la capacidad de identificar patrones en datos masivos y a la vez poder hacer predicciones (análisis predictivo). En la actualidad esta tecnología se encuentra en varias aplicaciones como las respuestas inteligentes de Gmail o el habla de Siri y Alexa. El machine Learning utiliza dos tipos de enfoques o técnicas:

- a. Aprendizaje Supervisado, es aquel que entrena un modelo sobre los datos de entrada y salida conocidos para que realice la predicción de resultados futuros.
- b. Aprendizaje No supervisado, que encuentra patrones ocultos o estructuras propias en los datos de entrada

El machine Learning funciona con 4 pasos básicos para crear una aplicación o modelo de machine Learning los cuales son realizados por los llamados científicos de datos que trabajan de forma estrecha con los negocios que desarrollan el modelo.

Paso 1. Seleccionar y preparar un conjunto de datos de entrenamiento.

Estos datos de entrenamiento son un conjunto de datos representativos de los datos que el modelo de Machine Learning que se introducirán para resolver el problema que está diseñado para resolver.

Paso 2. Elegir un algoritmo para ejecutarlo en el conjunto de datos de entrenamiento.

Este paso es el más importante se debe elegir qué algoritmo utilizar, siendo este un conjunto de pasos de procesamiento estadístico. El algoritmo a utilizar depende mucho del tipo con o sin etiqueta y la cantidad de datos de entrenamiento y del tipo de problema a resolver.

Entre estos tipos comunes de algoritmos de aprendizaje automático que se utilizan datos etiquetados se encuentran los siguientes.

- i. Algoritmos de regresión: la regresión lineal y logística
- ii. Árboles de decisión: son aquellos que utilizan datos clasificados para realizar recomendaciones basadas en un conjunto de reglas de decisión.
- iii. Algoritmos basados en instancias: son aquellos que se utilizan para realizar probabilidades de que un punto de datos sea miembro de un grupo u otro en función de su proximidad a otros puntos de datos.

Entre los algoritmos que se pueden utilizar con datos no etiquetados se encuentran los siguientes:

- i. Algoritmos de agrupación: La agrupación se enfoca en la identificación de grupos de registros similares y en el etiquetado de los registros según el grupo al que pertenecen

- ii. Algoritmos de asociación: los algoritmos de asociación encuentran patrones y relaciones en los datos e identifican relaciones frecuentes 'if-then' llamadas reglas de asociación. Estas son similares a las reglas usadas en la minería de datos.
- iii. Redes neuronales: es un algoritmo que define una red en capas de cálculos con una capa de entrada, donde se ingresan los datos; al menos una capa oculta, donde se realizan los cálculos y se sacan diferentes conclusiones sobre la entrada; y una capa de salida, donde a cada conclusión se le asigna una probabilidad.

Paso 3: Entrenamiento del algoritmo para crear el modelo

En este paso se realiza en proceso iterativo donde se ejecuta las variables a través del algoritmo comparando el resultado con los resultados que ha producido el algoritmo y tratando de ajustar el rumbo que podría dar un resultado más exacto realizando la ejecución de las variables hasta que el algoritmo devuelva el resultado correcto

Paso 4: Usar y mejorar el modelo

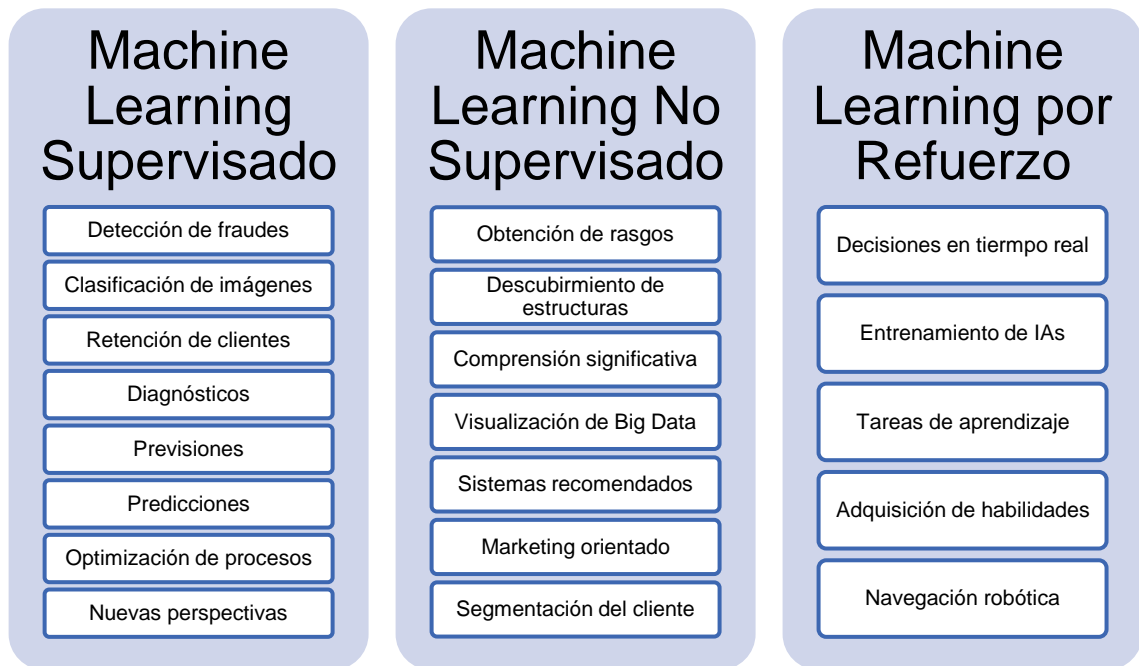
Este el paso final donde se realiza la utilización del modelo con los nuevos datos para que mejore la precisión y eficacia con el tiempo.

1.3.13. Deep Learning

Conocido como el aprendizaje profundo es un elemento muy importante en la ciencia de datos donde incluye estadísticas y modelos predictivos siendo muy beneficioso para los científicos de datos que tienen la tarea de recopilar, analizar e interpretar grandes cantidades de datos siendo que este proceso sea más rápido y sencillo se puede considerar como una forma de automatizar el análisis predictivo.

Figura 16

Aplicaciones del machine Learning



Nota. Fuente: Gonzáles et al. (2021).

Red feedforward

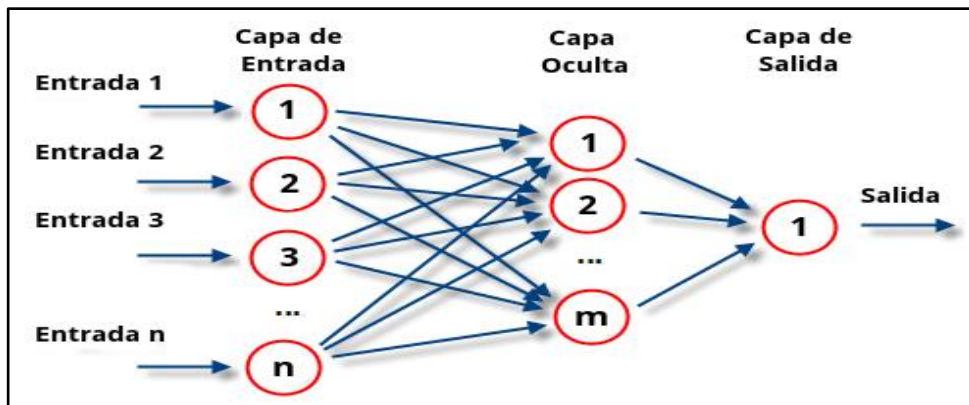
Es aquella red neuronal artificial conocida también como una red prealimentada donde las conexiones entre las unidades no forman un ciclo la topología del arreglo de neuronas y sus interconexiones hace fluir la información de forma unidireccional o una sola dirección para que nunca pueda pasar más de una vez a través de una neurona antes de generarse la respuesta de salida. Siendo diferentes a las redes neuronales recurrentes. Redes feedforward son aquellas que son útiles en aplicaciones de reconocimiento o clasificación de patrones. Las más conocidas son:

- i. Perceptrón
- ii. Adaline
- iii. Madaline
- iv. Backpropagation

En las Redes feedforward, cada círculo representa una neurona, cada neurona recibe señales desde las neuronas situadas a su izquierda y envía señales a las neuronas situadas a la derecha.

Figura 17

Modelo de red neuronal

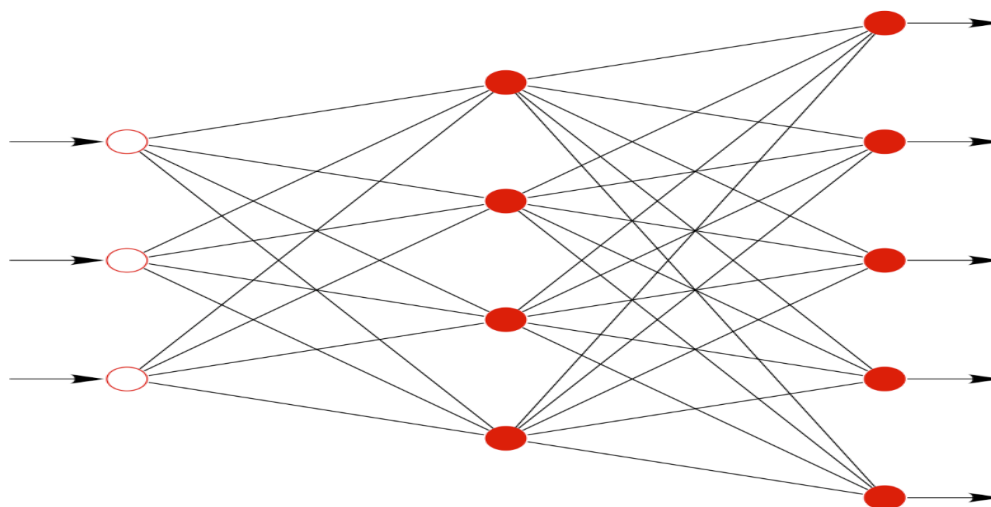


Nota. Fuente: ATRIA Innovation (2019).

El perceptrón multicapa es una red neuronal artificial (RNA) formada por múltiples capas, de tal manera que tiene capacidad para resolver problemas que no son linealmente separables.

Figura 18

Red neuronal Feedforward



Nota. Fuente: Salas (2016).

1.3.14. Algoritmo de Backpropagation

El algoritmo Backpropagation es uno de los algoritmos más usados en redes con conexión hacia adelante aplica la técnica de aprendizaje supervisado de gradiente descendente , está estructurado en dos fases :primero se aplica un patrón de entrada como estímulo para la primera capa de las neuronas el cual se propagaba a través de las distintas capas que componen la red hasta producir una salida, se compran el resultado de las neuronas de salida con la salida y se calcula un valor de error para cada neurona de salida y luego se transmiten hacia atrás partiendo de la capa de salida , hacia las neuronas de las capas intermedias este proceso se repite capa por capa hasta que todas las neuronas de la red hayan recibido un error que describa su contribución relativa al error total. Basándose en la señal de valor del error recibido, se reajustan o actualizan los pesos de conexión de cada neurona, de manera que en la siguiente vez que se presente el mismo patrón, la salida este más cercana a la deseada.

Lo importante de este proceso es que a medida que la red se entrena las neuronas de las capas intermedias se organizan así mismas de tal modo que las distintas neuronas se aprenden a reconocer cuando se les presente un patrón arbitrario de entrada que contenga ruido o que esté incompleto, las neuronas de la capa oculta de la red responderán con una salida activa si la nueva entrada contiene un patrón que se asemeje a aquella característica que las neuronas individuales hayan aprendido a reconocer durante su entrenamiento. Y a la inversa, las unidades de las capas ocultas tienen una tendencia a impedir su salida si el patrón de entrada no contiene la característica para reconocer, para la cual han sido entrenadas. Uno de los mayores problemas con el algoritmo BP ha sido el largo tiempo de entrenamiento debido a que el método de descenso por gradiente es sencillo, pero lento (Alpana, 2002).

1.3.15. Modelos de optimizadores

Según Keras (2021), se cuentan con los siguientes modelos de optimizadores:

1.3.15.1. Adam

La optimización de Adam es un método de descenso de gradiente estocástico que se basa en la estimación adaptativa de momentos de primer y segundo orden.

Según Kingma & Ba (2017), el método es “computacionalmente eficiente, tiene pocos requisitos de memoria, invariante al reajuste diagonal de gradientes y es muy adecuado para problemas que son grandes en términos de datos/parámetros”.

Figura 19

Optimizador que implementa el algoritmo de Adam

```
tf.keras.optimizers.Adam(  
    learning_rate=0.001,  
    beta_1=0.9,  
    beta_2=0.999,  
    epsilon=1e-07,  
    amsgrad=False,  
    name="Adam",  
    **kwargs  
)
```

Nota. Fuente, Keras (2021).

1.3.15.2. Adadelta

La optimización Adadelta es un método de descenso de gradiente estocástico que se basa en la tasa de aprendizaje adaptativo por dimensión para abordar dos inconvenientes:

- La continua decadencia de las tasas de aprendizaje a lo largo de la formación.
- La necesidad de una tasa de aprendizaje global seleccionada manualmente.

Adadelta es una extensión más sólida de Adagrad que adapta las tasas de aprendizaje en función de una ventana móvil de actualizaciones de gradientes, en lugar de acumular todos los gradientes anteriores. De esta manera, Adadelta continúa aprendiendo incluso cuando se han realizado muchas actualizaciones (Keras, 2021). En comparación con Adagrad, en la versión original de Adadelta no es necesario establecer una tasa de aprendizaje inicial. En esta versión, se puede establecer la tasa de aprendizaje inicial, como en la mayoría de los demás optimizadores de Keras.

Figura 20

Optimizador que implementa el algoritmo Adadelta

```
tf.keras.optimizers.Adadelta(  
    learning_rate=0.001, rho=0.95, epsilon=1e-07, name="Adadelta", **kwargs  
)
```

Nota. Fuente: Keras (2021).

1.3.15.3. Adagrad

Adagrad es un optimizador con tasas de aprendizaje específicas de parámetros, que se adaptan en relación con la frecuencia con la que se actualiza un parámetro durante el entrenamiento. Cuantas más actualizaciones reciba un parámetro, menores serán las actualizaciones.

Figura 21

Optimizador que implementa el algoritmo Adagrad

```
tf.keras.optimizers.Adagrad(  
    learning_rate=0.001,  
    initial_accumulator_value=0.1,  
    epsilon=1e-07,  
    name="Adagrad",  
    **kwargs  
)
```

Nota. Fuente: Keras (2021).

1.3.15.4. Adamax

Es una variante de Adam basada en la norma del infinito. Los parámetros predeterminados siguen los proporcionados en el documento. Adamax es a veces superior a Adam, especialmente en modelos con incrustaciones.

Figura 22

Optimizador que implementa el algoritmo Adamax

```
tf.keras.optimizers.Adamax(  
    learning_rate=0.001, beta_1=0.9, beta_2=0.999, epsilon=1e-07, name="Adamax", **kwargs  
)
```

Nota. Fuente: Keras (2021).

1.3.15.5. RMSprop

La esencia de RMSprop es:

- Mantener un promedio móvil (descontado) del cuadrado de gradientes
- Divida el gradiente por la raíz de este promedio.

Esta implementación de RMSprop utiliza un impulso simple, no el impulso de Nesterov.

La versión centrada además mantiene un promedio móvil de los gradientes y usa ese promedio para estimar la varianza.

Figura 23

Optimizador que implementa el algoritmo RMSprop

```
tf.keras.optimizers.RMSprop(  
    learning_rate=0.001,  
    rho=0.9,  
    momentum=0.0,  
    epsilon=1e-07,  
    centered=False,  
    name="RMSprop",  
    **kwargs  
)
```

Nota. Fuente: Keras (2021).

1.4. Formulación del Problema

¿Cómo lograr la predicción de ventas utilizando inteligencia de negocios y data analytics en la empresa Centro Textil de la Matta S.A.C.?

1.5. Justificación e importancia del estudio

En la investigación realizada apoyó a mejorar el proceso de ventas utilizando el método cuantitativo el cual le permitió obtener de sus datos de las ventas donde dicha información obtenida en su base de datos permitió realizar las proyecciones de ventas mediante machine learning con el objetivo que la empresa sea sostenible en el tiempo y ser competitiva en el mercado textil.

1.6. Hipótesis

Con el desarrollo de un modelo predictivo utilizando inteligencia de negocios y data analytics, se logrará la predicción de ventas en la empresa Centro Textil de la Matta S.A.C.

1.7. Objetivos

1.7.1. Objetivo general.

Desarrollar un modelo predictivo del proceso de ventas utilizando inteligencia de negocios y data analytics en la empresa Centro Textil de la Matta S.A.C.

1.7.2. Objetivos específicos.

- a. Seleccionar el modelo de aprendizaje automático más empleado para la predicción de ventas en la literatura científica.
- b. Diseñar el modelo predictivo para el proceso de ventas utilizando inteligencia de negocios y data analytics.
- c. Implementar el modelo predictivo del proceso de ventas en la empresa caso de estudio.
- d. Evaluar los resultados de la predicción de ventas utilizando el modelo predictivo propuesto.

II. MATERIAL Y MÉTODO

2.1. Tipo y Diseño de Investigación

Tipo de investigación.

Se empleó el tipo de investigación Tecnológica Aplicada puesto que, en el desarrollo del modelo predictivo se tuvieron que utilizar estrategias de IE, más específicamente BI y DA, para posteriormente ser aplicado en una empresa del sector textil limeño, con el que se lograron predecir las ventas de dicha empresa caso de estudio, la compañía Centro Textil de la Matta S.A.C.

Diseño de Investigación.

Se empleó el diseño de investigación Cuasi-Experimental, manipulando la variable dependiente “Predicción de Ventas”, la misma que fue evaluada en un diseño post prueba únicamente. Más específicamente, las ventas fueron pronosticadas en un escenario posterior a la puesta en marcha del modelo predictivo basado en BI y DA en la empresa Centro Textil de la Matta, haciendo un análisis de un dataset compuesto por las ventas realizadas durante los años 2017-2021.

Asimismo, el enfoque es cuantitativo porque se realizó un análisis predictivo de las ventas utilizando las informaciones respecto a las ventas históricas realizadas por la empresa Centro Textil de la Matta durante el periodo 2017-2021. Por ello, se realizó una aplicación del modelo predictivo basado en BI y DA, con el que se lograron resultados cuantitativos respecto a los indicadores mencionados en la Tabla 4.

2.2. Población y muestra

Población

La población estuvo constituida por la información con los datos de las ventas de la empresa Centro Textil de la Matta SAC durante los años 2017, 2018, 2019, 2020, 2021, las mismas que fueron utilizadas para poder entrenar el algoritmo predictivo correspondiente. Vale precisar que dichas informaciones han sido

datos raw, es decir, un dato sin ningún tipo de procesamiento, las cuales estuvieron compuestos por valor fecha (DD/MM/AA) y valor venta (VV). Obteniendo un total de 1460 registros.

Tabla 2

Población

N°	FECHA	VV
1	12/04/2019	7118.11
2	13/04/2019	4685.51
3	14/04/2019	1173.18
4	15/04/2019	12635.12
5	16/04/2019	13033.69
6	17/04/2019	26239.28
7	18/04/2019	4626.31
8	20/04/2019	979.56
9	21/04/2019	0.00
10	22/04/2019	6851.69
11	23/04/2019	10742.62
12	24/04/2019	18696.89
13	25/04/2019	22523.03
14	26/04/2019	15463.97
15	27/04/2019	12588.92
16	28/04/2019	927.94
17	29/04/2019	23318.75
18	30/04/2019	25866.18
19	2/05/2019	9264.63
20	3/05/2019	16902.80
21	4/05/2019	4069.27
22	5/05/2019	1368.66
23	6/05/2019	14949.57
24	7/05/2019	8507.23

Nota. Fuente: elaboración propia.

Muestra

La muestra estuvo constituida por los datos de las ventas de la empresa Centro Textil de la Matta SAC durante los años 2017, 2018, 2019, 2020, 2021, las mismas que fueron utilizadas para poder entrenar el algoritmo predictivo correspondiente. Vale precisar que dichas informaciones óptimas han sido datos limpios, es decir, en los cuales no existe error o nulidad, por ejemplo “NaN”, “0”, etcétera, las cuales estuvieron compuestos por valor fecha (DD/MM/AA) y valor venta (VV). Siendo un total de 1051 datos puros o limpios para ser utilizado en el proceso de predicción.

Tabla 3

Muestra

Nº	AÑO	DATA
1	2017	312
2	2018	316
3	2019	318
4	2020	279
5	2021	226
Total de datos (count)		1051

Nota. Fuente: elaboración propia.

2.3. Variables, Operacionalización

Variable independiente:

Modelo Predictivo utilizando Inteligencia de Negocios y Data Analytics.

Variable dependiente:

Predicción de ventas.

Tabla 4

Operacionalización de variables

Variables	Dimensiones	Indicadores	Ítem	Técnica e instrumentos de recolección de datos
Variable Independiente Modelo Predictivo utilizando Inteligencia de Negocios y Data Analytics	Tiempo de respuesta	Tiempo total del proceso	$T = \sum_n^e = \frac{1}{29} t_e$	T: Registro Electrónico
	Consumo de memoria	Memoria consumida durante un proceso	$C_m = \sum_j^n \frac{cm_j}{n}$	I: Ficha de Registro Electrónico
	Consumo de CPU	CPU consumido durante un proceso	$C_{CPU} = \sum_j^n \frac{cc_j}{n}$	Anexo 3
Variable Dependiente Predicción de ventas	Error	Error absoluto medio (MAE)	$MAE = \frac{\sum_{i=1}^n y_i - x_i }{n}$	T: Registro Electrónico
		Error cuadrático medio (MSE)	$MSE = \frac{\sum Error\ de\ pronóstico^2}{n}$	
		Error absoluto relativo (MAPE)	$MAPE = \frac{\sum_{i=1}^n Real_i - Pronóstico_i }{Real_i}$	I: Ficha de Registro Electrónico
	Exactitud	Desviación absoluta media (MAD)	$MAD = \frac{\sum_{t=1}^n V_r - \bar{V}_p }{n}$	Anexo 4

Nota. Fuente: elaboración propia.

2.4. Técnicas e instrumentos de recolección de datos, validez y confiabilidad

2.4.1. Técnicas de recolección de datos

Se emplearon las siguientes técnicas de recolección:

Recolección de datos

Se llevó a cabo la recolección de datos solicitando previamente el permiso respectivo a la Gerencia General de la empresa Centro Textil de la Matta SAC. Posterior a ello, se realizó un proceso de extracción de datos de las ventas realizadas durante el periodo 2017-2021 que se encontraba en la base de datos Postgre SQL en versión 2014 de nombre "MATTA" de dicha empresa caso de estudio. Para dicha actividad, se necesitó de un script realizado mediante lenguaje de programación Visual Fox-Pro para poder hacer la transacción de la BD SQL realizando la extracción de los datos del servidor hacia una hoja de cálculo MS-Excel, para luego ser procesada en Python.

Registro electrónico

Se llevó a cabo el registro electrónico de los indicadores especificados en la Tabla 4 de modo que se pudieran establecer datos cuantitativos reales acerca de los recursos utilizados por el modelo predictivo propuesto, el cual se basó en BI y DA. Asimismo, también se empleó este registro para cuantificar los errores y la exactitud del modelo propuesto en cuanto al pronóstico de ventas.

Entrevista

Se llevó a cabo la entrevista para establecer el estado previo a la puesta en marcha del modelo predictivo en la empresa Centro Textil de la Matta SAC, de modo que se pudieran conocer de primera mano cómo es que se ejecutaban los procesos de ventas, qué indicadores se empleaban y los procesos de manejo de información por parte del área de sistemas y administración respecto a las ventas en dicha empresa caso de estudio.

2.4.2. Instrumentos de recolección de datos

Se emplearon los siguientes instrumentos de recolección:

Ficha de recolección de datos

Se emplearon fichas de recolección para los datos históricos de las ventas de la empresa Centro Textil de la Matta SAC, los cuales, en una primera intervención consideraron varias características generales tal y como se muestra a continuación:

Tabla 5

Resumen de campos de las fichas de recolección

N°	Ítem	Descripción	Campo
1	Año	Indica el año de la venta	ANIO
2	Mes	Indica el mes de la venta	MES
3	Documento	Indica el tipo de documento de la venta	DOCU
4	Fecha	Indica la fecha de venta	FECHA
5	Cuenta cliente	Indica el código interno asignado a un cliente	CTACTE
6	Razón social	Indica la razón social del cliente	RAZON
7	Código	Indica el código del producto expendido	CODIGO
8	Descripción	Indica la descripción del producto	DESCRIP
9	Valor de venta	Indica el valor de venta del producto en soles	VV
10	Cantidad	Indica la cantidad de productos expendidos	CANT

Nota. Fuente: elaboración propia.

Ficha de registro electrónico

Se emplearon ficha de registro electrónico para cada uno de los indicadores especificados en la Tabla 4, tal y como se resume a continuación:

Tabla 6

Resumen de fichas de registro electrónico

N°	Indicador	Instrumento	Anexo
1	Tiempo total del proceso	Registro electrónico	Anexo 3
2	Memoria consumida	Registro electrónico	Anexo 3
3	CPU consumido	Registro electrónico	Anexo 3
4	MAE	Registro electrónico	Anexo 4
5	MSE	Registro electrónico	Anexo 4
6	MAPE	Registro electrónico	Anexo 4
7	MAD	Registro electrónico	Anexo 4

Nota. Fuente: elaboración propia.

Guía de entrevista

Se emplearon tres (03) fichas de entrevista para conocer la situación previa a la aplicación del modelo en el caso de estudio, la empresa Centro Textil de la Matta SAC, de modo que se obtuviera información acerca de la problemática acaecida en dicha empresa en cuanto al pronóstico de las ventas. Dichas fichas se encuentran resumidas a continuación:

Tabla 7

Resumen de guías de entrevista

N°	Área	Entrevistado	Anexo
1	Gerencia General	Jaime Matta Coronado	Anexo 14
2	Administrador	Carlos Obando Tuesta	Anexo 14
3	Encargado de sistemas	Luis Paredes Rosario	Anexo 14

Nota. Fuente: elaboración propia.

2.5. Procedimiento de análisis de datos

2.5.1. Variable Independiente

Para lograr la evaluación de la Variable Independiente, más específicamente, para el Modelo Predictivo utilizando BI y DA, se utilizaron los siguientes indicadores:

En la siguiente ecuación, el indicador “Tiempo total del proceso” es el tiempo empleado por época y sus valores son entre 1s y 1ms.

$$T = \sum_{n=29}^{e=1} t_e$$

Donde:

T= Tiempo Total del proceso

e = Época

t = Tiempo empleado por época

En la siguiente ecuación, el indicador “Memoria consumida durante un proceso” es la cantidad de memoria RAM que es empleada durante el proceso, por ejemplo, de entrenamiento de la red neuronal.

$$C_m = \sum_j^n \frac{cm_j}{n}$$

Donde:

C_m = Es el grado de consumo de memoria del clasificador

cm_j = Es el grado de consumo de memoria en la prueba j

n : Es el total de pruebas

En la siguiente ecuación, el indicador “CPU consumido durante un proceso” es la cantidad de CPU que es empleado durante el proceso, por ejemplo, de entrenamiento de la red neuronal.

$$C_{CPU} = \sum_j^n \frac{cc_j}{n}$$

Donde:

C_{CPU} = Es el grado de consumo de CPU del clasificador

cc_j = Es el grado de consumo de CPU en la prueba j

n = Es el total de pruebas

2.5.2. Variable Dependiente

Para lograr la evaluación de la Variable Dependiente, más específicamente, para la Predicción de Ventas, se utilizaron los siguientes indicadores:

En la siguiente ecuación, el indicador “Error absoluto medio (MAE)” es el resultado que se obtiene de la diferencia entre el valor pronosticado y el valor real en cada punto que se pronostica, en este caso de cada época.

$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n}$$

Donde:

MAE = Error absoluto medio

y_i = Predicción del modelo

x_i = Valor Verdadero

n = épocas

En la siguiente ecuación, el indicador “Error cuadrático medio (MSE)” el promedio de los cuadrados de las diferencias entre los valores pronosticados y los reales de nuestro modelo de nuestra red neuronal definiendo la interacción a 70 épocas en nuestro test de entrenamiento.

$$MSE = \frac{\sum Error\ de\ pronóstico^2}{n}$$

Donde:

MSE = Error cuadrático medio

n = épocas

En la siguiente ecuación, el indicador “Error absoluto relativo (MAPE)” el promedio de las diferencias absolutas encontradas entre los valores pronosticados y los reales el error nos da a como un porcentaje de los valores reales de nuestro modelo.

$$MAPE = \frac{\sum_{i=1}^n |Real_i - Pronóstico_i|}{\frac{Real_i}{n}}$$

Donde:

$MAPE$ = Error absoluto relativo

$Real_i$ = Valores reales del error

$Pronóstico_i$ = Valores reales del pronóstico

n = épocas

En la siguiente ecuación, el indicador “Desviación absoluta media” es la sumatoria total de diferencia entre las ventas realizadas en un periodo de tiempo determinado menos el valor absoluto del promedio de dichas ventas realizadas, dividido entre el número total del periodo de tiempo específico.

$$MAD = \frac{\sum_{t=1}^n |V_r - \bar{V}_p|}{n}$$

Donde:

MAD = Desviación absoluta media

V_r = Valor de ventas realizadas en un periodo de tiempo

V_p = Valor de ventas promedio realizadas en periodo de tiempo

n = periodo de tiempo

2.6. Criterios éticos

Se emplearon los siguientes criterios éticos:

- a. Confidencialidad: “Las informaciones de las ventas de la empresa Centro Textil de la Matta S.A.C. han sido tratadas en estricta reserva ya que eran datos privados de dicha compañía, por lo que el autor no ha hecho pública el listado de dichas ventas” (Noreña et al., 2012).

- b. Derechos de autor: “Se emplearon citas para hacer referencia a cada una de las fuentes que sirvieron de soporte académico para el conocimiento científico necesario en esta investigación, las cuales fueron posteriormente localizadas en las referencias de esta investigación” (Noreña et al., 2012).
- c. Originalidad: “Se estableció el nivel de originalidad de la investigación haciendo uso del software anti plagio Turnitin con lo que se obtuvo un nivel porcentual máximo del 15% de similitud con otras fuentes” (Erazo, 2011).

2.7. Criterios de Rigor Científico

Se emplearon los siguientes criterios de rigor científico:

- a. Confiabilidad: “El modelo predictivo basado en BI y DA pronosticó el proceso de ventas considerando indicadores para la fiabilidad de los resultados, tales como, verbigracia, MAE, MSE, MAPE y la precisión del modelo mismo” (Hofseth, 2018).
- b. Autenticidad: “La investigación fue auténtica ya que partió de una problemática social y de ingeniería que involucraba a las empresas de la actividad manufacturera, más específicamente, a las empresas del sector textil del Emporio Comercial Gamarra, por lo que el modelo basado en BI y DA pretendió pronosticar las ventas de dichas microempresas peruanas” (Arias & Giraldo, 2011).
- c. Veracidad: “Los resultados obtenidos posteriores a la puesta en marcha del modelo predictivo del proceso de ventas utilizando BI y DA fueron veraces, de tal manera que han podido ser cotejadas con otras investigaciones de otros autores, logrando discutir ambos resultados, para contrastarlos o contraponerlos” (Arias & Giraldo, 2011).

III. RESULTADOS

3.1. Resultados en tablas y figuras

3.1.1. Resultado de la Variable Independiente: Modelo Predictivo utilizando Inteligencia de Negocios y Data Analytics

Para ejecutar la predicción de 7 días respecto a las ventas en la empresa caso de estudio, se utilizaron un total de 1460 registros, de los cuales, se hizo el proceso de limpieza, obteniendo un total de 1051 registros óptimos para luego ser estructurados de modo que, el 80% se empleó para el entrenamiento (training) y el 20% restante se empleó para el testeo (testing), para que, de este modo, se pueda ejecutar con mayor eficiencia la red neuronal. Luego, se consideraron los datos históricos de las ventas realizadas en un total de 29 días, los mismos que sirvieron, a posteriori, para la predicción de las ventas de los 7 días posteriores a la fecha fin. Para dicho proceso se utilizó la red feedforward y se consideraron los indicadores mencionados en la Tabla 4 teniendo en cuenta las variables delimitadas.

3.1.1.1. Indicador “Tiempo total del proceso”

$$T = \sum_{n=29}^{e=1} t_e$$

$$T = 134ms$$

Interpretación. Se obtuvo un resultado para el indicador “Tiempo total del proceso” de 134ms considerando para ello, un total de 29 días, partiendo de una fecha inicial de “2021-08-20” hasta “2021-09-23” para que la red neuronal pueda aprender a predecir las ventas futuras.

3.1.1.2. Indicador “Memoria consumida durante un proceso”

Para obtener la memoria consumida en el proceso de entrenamiento del modelo se aplicó el uso de la librería psutil de Python. Donde se realizó lo siguiente:

- Se consideraron los valores obtenidos antes y después de uso de memoria en el proceso de entrenamiento.
- Se consideraron un total de 29 días, partiendo de una fecha inicial de “2021-08-20” hasta “2021-09-23” para el proceso de entrenamiento de la red neuronal.
- Valor de inicio de memoria antes de ejecución del proceso de entrenamiento. Valor de 0.0 Mb y el pico de 0.0 Mb
- Valor de termino después de término del proceso entrenamiento. Valor 0.55453 MB y el pico fue de 0.560679 MB.
- Obteniendo el total de uso de Memoria 0.55453 MB.

$$C_m = \sum_j^n \frac{cm_j}{n}$$

$$C_m = 0.55453 \text{ MB.}$$

3.1.1.3. Indicador “CPU consumido durante un proceso”

En este indicador se utilizó la aplicación de la librería psutil de Python determinando el inicio y termino del uso de la CPU. Este proceso se realizó de la siguiente manera.

- Se consideraron un total de 29 días, partiendo de una fecha inicial de “2021-08-20” hasta “2021-09-23” para el proceso de entrenamiento de la red neuronal.
- Obteniendo el total de uso de CPU 8.3 GB

$$C_{CPU} = \sum_j^n \frac{CC_j}{n}$$

$$C_{CPU} = 8.3 \text{ GB.}$$

3.1.2. Resultados de la Variable Dependiente: Predicción de ventas

3.1.2.1. Indicador “Error absoluto medio (MAE)”

El indicador MAE es el resultado que se obtiene de la diferencia entre el valor pronosticado y el valor real que se pronostica en este caso. Para este caso específico, se consideró en el modelo de esta red neuronal la interacción a 29 días para evaluar la efectividad del modelo, obteniéndose como resultado el siguiente:

$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n}$$

$$MAE = 0.31008386611938477$$

3.1.2.2. Indicador “Error cuadrático medio (MSE)”

Este Indicador dio a conocer el promedio de los cuadrados de las diferencias entre los valores pronosticados y los reales del modelo de la red neuronal. Para este caso específico, se consideró en el modelo de esta red neuronal la interacción a 29 días para evaluar la efectividad del modelo, obteniéndose como resultado el siguiente:

$$MSE = \frac{\sum Error \text{ de pronóstico}^2}{n}$$

$$MSE = 0.1358100026845932$$

3.1.2.3. Indicador “Error absoluto relativo (MAPE)”

En este indicador dio a conocer como el promedio de las diferencias absolutas encontradas entre los valores pronosticados y los reales el error da como un porcentaje de los valores reales del modelo. Para este caso específico, se consideró en el modelo de esta red neuronal la interacción a 29 días para evaluar la efectividad del modelo, obteniéndose como resultado el siguiente:

$$MAPE = \frac{\sum_{i=1}^n |Real_i - Pronóstico_i|}{Real_i \cdot n}$$

$$MAPE = 6.578026294708252$$

3.1.2.4. Indicador “Desviación absoluta media (MAD)”

En este indicador da a conocer la desviación absoluta media del proceso de entrenamiento de test de entrenamiento siendo una medida de uso común para determinar el error de pronósticos la desviación absoluta promedio o MAD. Donde se suman las desviaciones absolutas y se obtiene un promedio. Para este caso específico, se consideró en el modelo de esta red neuronal la interacción a 29 días para evaluar la efectividad del modelo, obteniéndose como resultado el siguiente:

$$MAD = \frac{\sum_{t=1}^n |V_r - \bar{V}_p|}{n}$$

$$MAD_{real} = 0.1748294085264206$$

$$MAD_{pred} = 0.21287952363491058$$

Tabla 8*Resumen resultante de los pesos de los 29 días*

N°	MAD_{real}	MAD_{pred}	MSE	MAE	MAPE
1	-0.42406392	-0.48075941	0.00321438	0.05669549	0.13369563
2	-0.51942843	-0.50349236	0.00025396	0.01593607	0.03068002
3	-0.5507864	-0.54907107	2.94E-06	0.00171533	0.00311432
4	-0.505246	-0.5496586	0.00197248	0.0444126	0.08790292
5	-0.52946615	-0.50327277	0.00068609	0.02619338	0.0494713
6	-0.53166866	-0.53527135	1.30E-05	0.00360269	0.00677619
7	-0.42444962	-0.54475296	0.01447289	0.12030334	0.28343372
8	-0.4495166	-0.47235426	0.00052156	0.02283766	0.05080494
9	-0.22416633	-0.51407623	0.08404775	0.2899099	1.29328032
10	-0.33389324	-0.28756082	0.00214669	0.04633242	0.13876417
11	-0.22692794	-0.49449176	0.0715904	0.26756382	1.17906951
12	-0.08673579	-0.28676325	0.04001099	0.20002747	2.30617001
13	-0.40658486	-0.24970792	0.02461037	0.15687694	0.38584058
14	-0.77711886	-0.49230072	0.08112137	0.28481814	0.36650525
15	-0.61240107	-0.64112365	0.00082499	0.02872258	0.04690159
16	-0.23112929	-0.47793022	0.0609107	0.24680093	1.06780464
17	-0.19674432	-0.15346704	0.00187292	0.04327728	0.21996709
18	-0.49802357	-0.31280464	0.03430605	0.18521893	0.37190796
19	-0.31825012	-0.61239523	0.08652134	0.29414511	0.92425765
20	-0.31792545	-0.39861256	0.00651041	0.08068711	0.25379254
21	-0.44284636	-0.3946127	0.00232649	0.04823366	0.10891737
22	-0.59144074	-0.43527073	0.02438907	0.15617001	0.26405015
23	-0.6149707	-0.6068064	6.67E-05	0.0081643	0.01327592
24	-0.23731261	-0.54609793	0.09534838	0.30878532	1.30117537
25	-0.14968115	-0.20023005	0.00255519	0.0505489	0.33771051
26	-0.4544654	-0.30110955	0.02351802	0.15335585	0.3374423
27	-0.38706946	-0.56870753	0.03299239	0.18163807	0.46926478
28	-0.28676414	-0.46344557	0.03121633	0.17668143	0.61612109
29	-0.4687273	-0.3462978	0.01498898	0.1224295	0.26119558

Nota. Fuente, elaboración propia.

3.2. Discusión de resultados

Respecto al objetivo, Seleccionar el modelo de aprendizaje automático más empleado para la predicción de ventas en la literatura científica, se realizó un método de revisión sistemática de la literatura constituido por cinco (05) fases con el propósito de determinar tal modelo con lo cual se obtuvo como resultado al modelo MLP-Backpropagation como el de mayor uso de entre dichos modelos encontrados, con un total de tres (03) estudios (30%), entre los cuales se tuvieron, Regresión Logística, Support Vector Machine y Naive Bayes. Estos resultados concuerdan con los obtenidos por Wang et al. (2019) quienes construyeron un modelo basado en MLP-Backpropagation con el cual obtuvieron como resultado del error cuadrático medio (RMSE) de predicción es del 5,0%, lo que tiene un buen valor de aplicación determinada, con lo que concluyeron que, la red neuronal de BP es el modelo de predicción que puede predecir con precisión el volumen de ventas de vehículos de nueva energía.

Respecto al objetivo, Diseñar el modelo predictivo para el proceso de ventas utilizando inteligencia de negocios y data Analytics, se diseñó dicho modelo considerando cuatro (04) procesos principales: a. Proceso de Ingreso; b. Proceso de Análisis; c. Proceso de Ejecución y; d. Proceso de Salida con el que se pudiera predecir las ventas de la empresa caso de estudio relacionado de modo que se lograran identificar patrones en los comportamientos de clientes y clientes potenciales para realizar predicciones de ventas futuras en un periodo determinado siete días. Estos resultados se contraponen con los obtenidos por Dairu & Shilong (2021) quienes también desarrollaron un modelo para la predicción de ventas pero que, sin embargo, determinaron dos áreas, la clasificación y regresión para construir el conocimiento de las ventas de mercado a partir de datos estadísticos de ventas durante los últimos 1913 días de las tiendas Walmart en tres estados y pronosticar las ventas en los próximos 28 días, es decir cuatro semanas, mientras que este estudio predijo las ventas de 1 semana, y para ello utilizaron XGBoost. En comparación con otros métodos de aprendizaje en conjunto, XGBoost se ejecuta diez veces más rápido

utilizando mucho menos recursos, es sensible a los valores atípicos, primero hacemos un preprocesamiento de datos para filtrar valores atípicos, luego se convierten en tipos de datos flotantes para guardar en memorial.

Respecto al objetivo, Implementar el modelo predictivo del proceso de ventas en la empresa caso de estudio, se implementó dicho modelo antes mencionado en la empresa Centro Textil de la Matta S.A.C., teniendo en cuenta los datos de las ventas diarias (de los últimos 5 años), para pronosticar las ventas de cada día tomando en consideración los datos históricos de los 7 días anteriores. Para poder predecir las ventas diarias, se realizó un análisis de regresión de series de tiempo utilizando información de desplazamiento histórica combinada con una red neuronal de siete capas de entrada siete capas ocultas y una capa de salida. Estos resultados se contraponen con los obtenidos por Giri et al. (2019) quienes su arquitectura la basaron utilizando un modelo de inicio de red neuronal profunda V3 para extraer los vectores de características de las imágenes. Cada imagen definió un elemento único y se llevó a cabo en imágenes de ropa y sus ventas. Los datos de información fueron de 320 imágenes de ropa de mujer. Antes de fusionar los datos se dividió en entrenamiento y prueba considerando el 90% de los ítems para la formación y el 10% de los elementos para la prueba, es decir, 289 imágenes para entrenamiento y 31 imágenes para prueba. Después de fusionar la función de vectores de imágenes de productos con datos de ventas, el número final de observaciones fue 7403 y 748 en datos de entrenamiento y prueba.

Respecto al objetivo, Evaluar los resultados de la predicción de ventas utilizando el modelo predictivo propuesto se logró determinar que, Para poder obtener la predicción de los 7 días a partir del día 24/09/2021 al 30/09/2021 se analizó los datos reales de los días 16/09/2021 al 23/09/2021, para ver la tendencia de los valores de nuestro modelo al brindar los pronósticos de próximos 7 días verificándose que, existe una tendencia optima sin mostrar decadencia en los valores de venta en los próximos días. Estos resultados se contrastan con los

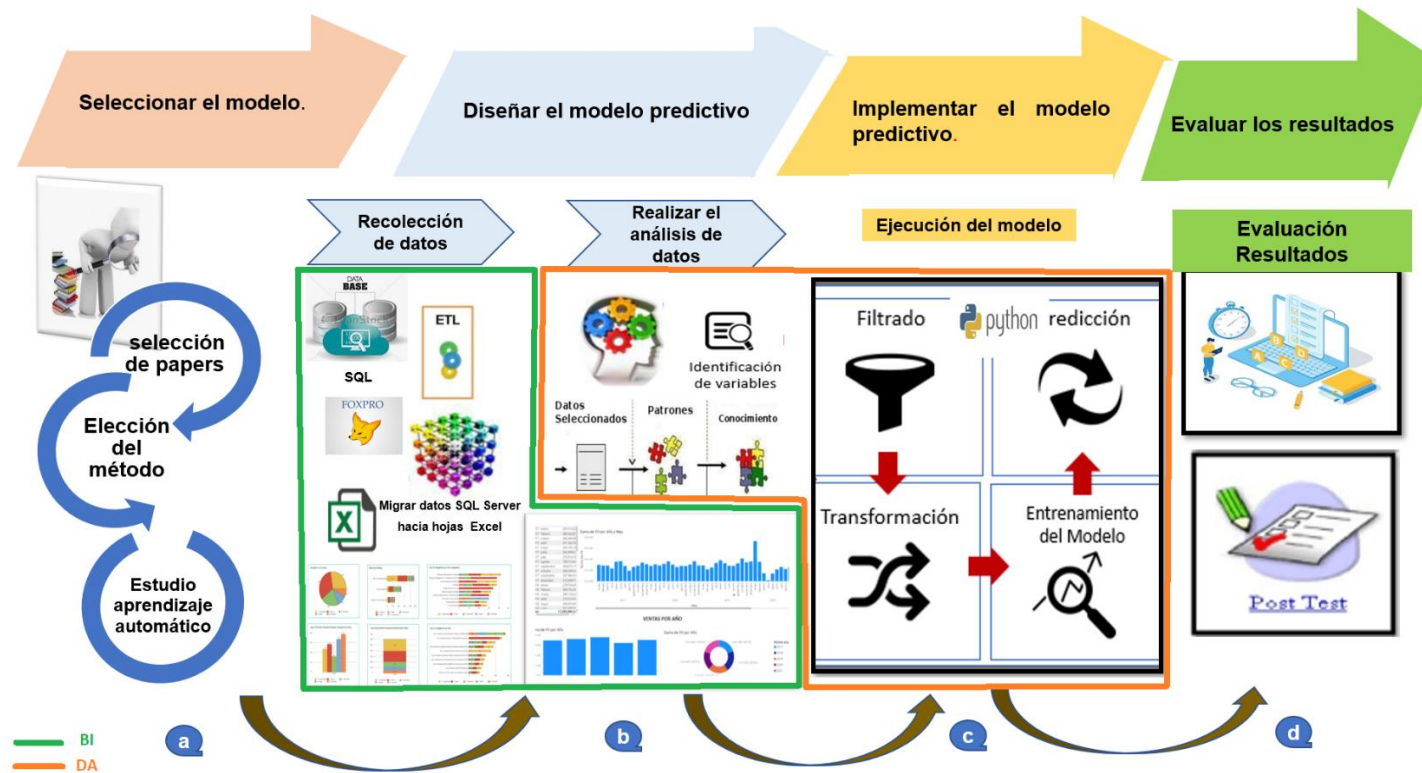
obtenidos por Punam et al. (2018) quienes establecieron el uso de un modelo estadístico de dos niveles como el más preciso obteniendo la media absoluta con un valor de MAE 0.3917. Siendo el modelo que superó las otras técnicas predictivas y contribuyó con mejores predicciones al conjunto de datos de Big Mart.

3.3. Aporte práctico

Para el desarrollo del siguiente informe de investigación se desplegó el método siguiente:

Figura 24

Método para el desarrollo del informe de investigación



Nota. Fuente, elaboración propia.

3.3.1. Modelos de aprendizaje automático para la predicción de series temporales de ventas

3.3.1.1. Acerca de los modelos de predicción de ventas

Obtener una predicción de ventas precisa es casi tan importante como alcanzar el objetivo de ingresos en sí. Según CSO Insights, citado por Pici (2021) “el 60% de los acuerdos previstos no se cierran realmente”. Como era de esperar, los datos también muestran que el 25% de los gerentes de ventas no están contentos con la precisión de sus pronósticos. La elección de un modelo de predicción adecuado puede marcar una gran diferencia en la capacidad para predecir con precisión los ingresos futuros.

Según Dairu & Shilong (2021) existen numerosos modelos que permiten la predicción de ventas tales como el modelo de Regresión Lineal, Support Vector Machines, el modelo Naive Bayes, etcétera, los cuales dan muy buenos resultados en cuanto a diversas problemáticas relacionadas con dichas predicciones de series temporales de ventas, tal y como puede visualizarse a continuación en la siguiente:

Tabla 9*Estudio acerca de modelos de aprendizaje automático*

Autor	Modelo	Descripción
Wang et al. (2021)	Regresión Logística	Se introdujo la teoría de la regresión logística y se propuso un método de detección de vehículos basado en la regresión logística.
Wong & Yeh (2019)	Random Forest	Se introduce un modelo de aprendizaje de Random Forest para ampliar el ámbito de aplicación del divide y vencerás, reduciendo así de forma fiable el tamaño del espacio de búsqueda.
Chen et al. (2018)	Modelo Bayesiano	El documento presentó principalmente la clasificación bayesiana como uno de los modelos de procesamiento de datos importantes en el campo del aprendizaje automático y la minería de datos.

Nota. Fuente, Dong et al. (2021)

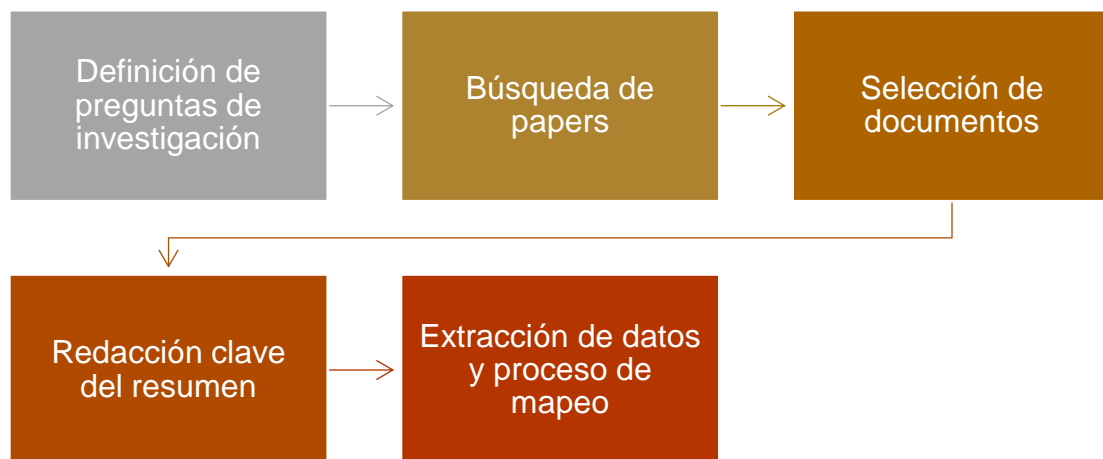
Por tanto, en esta parte de la presente investigación, se pretendió seleccionar el modelo de aprendizaje automático adecuado para la predicción de ventas, de modo que permita determinar el modelo más idóneo para dicho fin, el cual es el propósito general.

3.3.1.2. Elección del método de revisión sistemática de la literatura

Se llevó a cabo una revisión sistemática en cuanto a artículos basado en el método formulado por Petersen et al. (2008) el cual consta de cinco (05) pasos: (1) definición de las preguntas de investigación, (2) búsqueda de artículos, (3) selección de documentos, (4) redacción clave del resumen, (5) extracción de datos y proceso de mapeo. A continuación, se muestra dicho método de manera gráfica:

Figura 25

Pasos del método de revisión de la literatura utilizado



Nota. Fuente: Petersen et al. (2008).

3.3.1.3. Definición de preguntas de investigación

Este paso tuvo como propósito encontrar la información de la investigación existente acerca de los modelos de aprendizaje automático respecto a la predicción del proceso de ventas en diversas organizaciones. Para formular la pregunta de investigación, se siguió la estructura PICOC. Según Bruzza et al. (2017) PICOC es empleado a manera de método para describir los cinco (05) elementos de una pregunta de búsqueda. La estructura de dicho método es descrita a continuación:

Tabla 10*Estructura del Método PICOC*

P	I	C	O	C
Población	Intervención	Comparación	Resultados	Contexto
Individuo o grupo con una condición en particular	Manejo o intervención principal de interés	Hace alusión a una alternativa principal con la cual comparar	Lo que está tratando de lograr, medir, mejorar o afectar	Tipo de organización y bajo qué circunstancia.

Nota. Fuente, Bruzza et al. (2017).

A continuación, se muestra la estructura PICOC delimitada para la siguiente investigación:

Tabla 11*Estructura PICOC de la investigación*

PICOC	Descripción	Fórmula
P	El objetivo de la investigación	Modelo predictivo
I	Especificar aspectos de búsqueda o temas de interés para los investigadores	predicción de ventas, pronóstico de ventas, pronóstico de series de tiempo de ventas
C	Aspecto de la investigación con la que se compara la intervención	No aplica comparación
O	El escenario de la intervención	Procesos de ventas
C	Contexto en el que se investigará	Inteligencia de negocio, data analytics

Nota. Fuente, elaboración propia

A continuación, se muestran las preguntas de investigación empleadas para la siguiente investigación:

Tabla 12

Preguntas de investigación

N°	Pregunta de investigación	Descripción
P1	¿Cuántos artículos hablan del desarrollo de un modelo predictivo para el proceso de ventas por año?	Identificar la distribución anual de los trabajos que aplican un modelo predictivo para el proceso de ventas.
P2	¿Cuántos tipos de documentos hablan del desarrollo de un modelo predictivo para el proceso de ventas?	Identificar la distribución del tipo de paper que desarrolla un modelo predictivo para el proceso de ventas.
P3	¿Qué tipo de modelos predictivos existen para el proceso de ventas?	Identificar el modelo más utilizado.

Nota. Fuente, elaboración propia.

3.3.1.4. Búsqueda de artículos científicos

El segundo paso del proceso de mapeo sistemático fue buscar los artículos de revistas y actas de conferencia relacionados con las preguntas de investigación anteriormente mencionadas. La estructura PICOC es una buena forma de crear la cadena de búsqueda y debe estar impulsada por las preguntas de investigación. Puede tomarse de cada aspecto de la estructura. De la estructura PICOC en la Tabla 11, se evaluaron algunas de las palabras clave utilizadas para filtrar los papers, tales como: “Machine Learning Model”, “Sales Forecast Model”; “Sales Forecasting”; “Sales Prediction”; “Sales Time Series Forecasting”; “Efficient Sales Forecasting Model”; “Accurate Sales Forecasting Model”; “Exact Sales Predication”, utilizando los operadores booleanos “AND”, “OR”.

3.3.1.5. Selección de documentos

La cadena de búsqueda encuentra varios tipos de artículos en los temas. Aquí, podemos fusionar todos los datos de los documentos y elegir el artículo científico según los criterios relevantes. Para esto, se utilizaron bases de datos puestas a disposición por la casa de estudios a la que pertenece el autor y las mismas que se encuentran descritas a continuación:

Tabla 13

Fuentes de información empleadas

N°	Fuente de información	Página web
1	EBSCO	https://www.ebsco.com/
2	IEEE Xplore	https://ieeexplore.ieee.org/
3	ProQuest	https://www.proquest.com/
4	SCOPUS	https://www.scopus.com/
5	Science Direct	https://sciencedirect.com/

Nota. Fuente, elaboración propia.

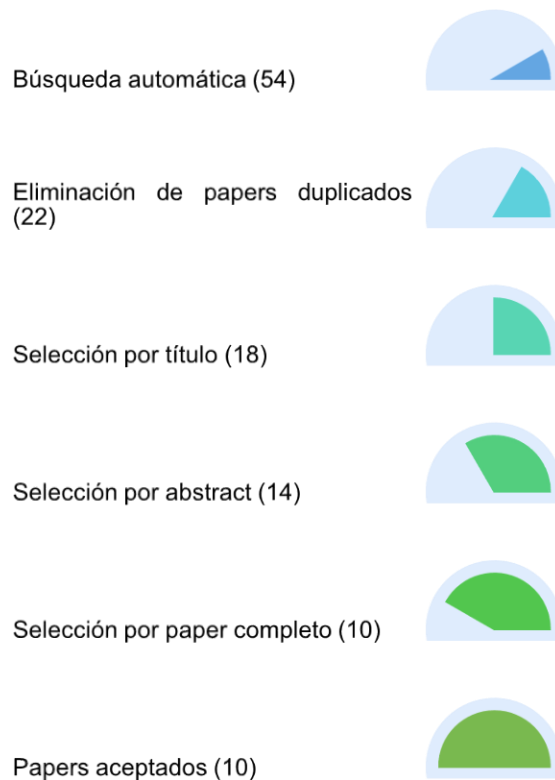
3.3.1.6. Desarrollo del método de selección de papers

En dichas bases de datos, se pudieron buscar los artículos escribiendo una cadena de búsqueda en una casilla de verificación de título, resumen y texto de contenido completo. Después de recuperar los resultados, se aplicaron criterios de selección para filtrar los candidatos.

A continuación, se muestra el resultado del proceso de selección, que consta de seis (06) pasos:

Figura 26

Resultado del proceso de selección



Nota. Fuente, elaboración propia.

A continuación, se muestran los seis (06) pasos seguidos en el proceso de selección de papers para la presente investigación:

1. Búsqueda automática: el resultado basado en las cadenas de búsqueda derivadas de las preguntas de investigación mostró cincuenta y cuatro (54) artículos que incluyeron revistas y actas de conferencias.
2. Eliminación de papers duplicados: se pudieron encontrar varios artículos en las cinco bases de datos mencionadas previamente, lo que conllevó a obtener resultados duplicados. Por lo tanto, se eliminaron los mismos

papers para evitar la redundancia de datos, y se identificaron veintidós (22) artículos independientes.

3. Selección por título: El objetivo de este filtro era eliminar documentos con un alcance claramente ajeno a la predicción de ventas. Por tanto, se eliminaron cuatro papers dado que ofrecían pronósticos en otras temáticas ajenas a la de ventas. Por lo tanto, posterior a la depuración se identificaron dieciocho (18) artículos científicos.
4. Selección por abstract: En esta etapa, se descartaron todos aquellos trabajos que no presentaban la implementación del modelo predictivo de ventas. Se eliminaron cuatro papers que no eran relevantes porque llegaban hasta el diseño del modelo, mas no a la implementación del mismo, identificándose catorce (14) papers.
5. Selección por paper completo: En el siguiente proceso, se retiraron los trabajos que no cumplían correctamente con la medición de los indicadores de predicción, por lo que se eliminaron cuatro papers que no eran relevantes. Por tanto, se distinguieron diez (10) papers.
6. Papers aceptados: Finalmente, de la lista inicial de cincuenta y dos (52) papers se consideraron solamente diez (10) de ellos los mismos que, fueron seleccionados para la evaluación y análisis.

Tabla 14*Crterios de inclusión y exclusión*

Criterio	Descripción
Inclusión	Investigación centrada en predicción de ventas. En la investigación académica a gran y pequeña escala Publicaciones internacionales. Publicaciones desarrolladas entre los años 2016-2021.
Exclusión	No asociado con la pregunta de investigación. El paper no pertenece a una revista o actas de conferencia.

Nota. Fuente, elaboración propia.

Por tanto, luego de haber aplicado el método propuesto por Petersen et al. (2008) se obtuvieron los siguientes resultados:

Tabla 15*Papers obtenidos de la revisión bibliográfica*

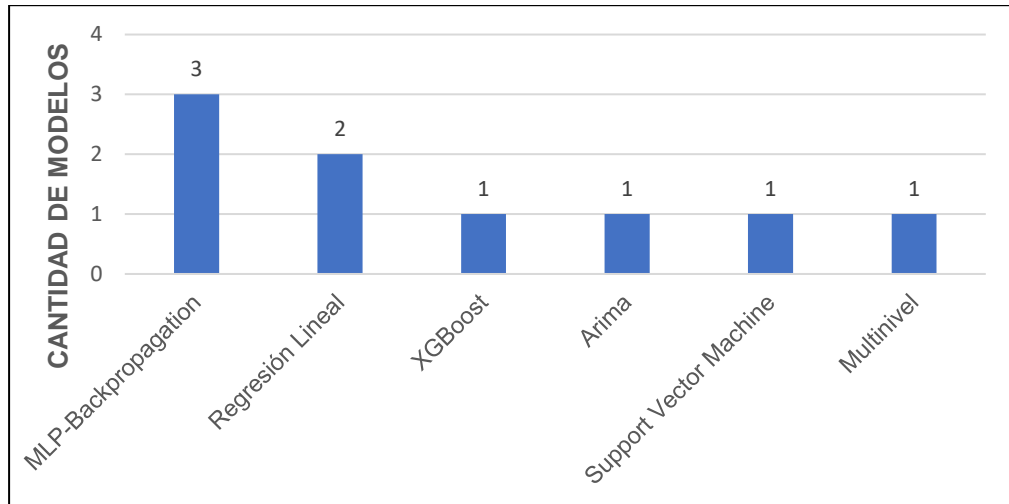
N°	Año	Autor	Título	Tipo de artículo	Modelo predictivo utilizado
1	2021	Dairu & Shilong	Machine Learning Model for Sales Forecasting by Using XGBoost	Acta de conferencia	Modelo XGBoost
2	2021	Dong et al.	A Machine Learning Model for Product Fraud Detection Based On SVM	Acta de conferencia	Modelo Support Vector Machine
3	2021	Yu Fan et al.	Prediction of A-share trend based on machine learning technologies	Acta de conferencia	Modelo de la red neuronal Back Propagation
4	2020	Bowen et al.	Forecasting method of e-commerce cargo sales based on ARIMA-BP model	Acta de conferencia	Modelo ARIMA
5	2020	Pavlyshenko	Machine-Learning Models for Sales Time Series Forecasting	Acta de conferencia	Modelo de regresión lineal
6	2019	Abdellatif et al.	Egyptian Case Study-Sales forecasting model for automotive section	Acta de conferencia	Modelo de regresión lineal
7	2019	Wang et al.	Sales Forecast of Chinese New Energy Vehicles Based on Wavelet and BP Neural Network	Acta de conferencia	Modelo de la red neuronal Back Propagation
8	2019	Giri et al.	Forecasting New Apparel Sales Using Deep Learning and Nonlinear Neural Network Regression	Acta de conferencia	Modelo de la red neuronal Back Propagation
9	2018	Punam et al.	A Two-Level Statistical Model for Big Mart Sales Prediction	Acta de conferencia	Modelo Multinivel
10	2018	Zheng & Li	Unmanned Retail's Distribution Strategy Based on Sales Forecasting	Acta de conferencia	Modelo de regresión lineal

Nota. Fuente, elaboración propia.

Por tanto, se ha podido distinguir que el Modelo de Back Propagation es el modelo predictivo más empleado y adecuado para el proceso de predicción de ventas según la literatura científica, tal y como puede visualizarse a continuación:

Figura 27

Modelos utilizados en la predicción de ventas



Nota. Fuente, elaboración propia.

Por tanto, se logró determinar que para esta investigación se utilizó el criterio del modelo de mayor usanza en la literatura científica, siendo el modelo de aprendizaje automático más adecuado para la predicción de ventas. el Modelo de Back Propagation, tal y como se muestra a continuación:

Tabla 16

Modelos más utilizados en la literatura científica

Modelo	# papers	%
MLP-Backpropagation	03	30%
Regresión Logística	02	20%
Support Vector Machine	01	10%
XGBoost	01	10%
ARIMA	01	10%

SVM	01	10%
Multinivel	01	10%

Nota. Fuente, elaboración propia.

3.3.2. Modelo predictivo para el proceso de ventas utilizando inteligencia de negocios y data analytics

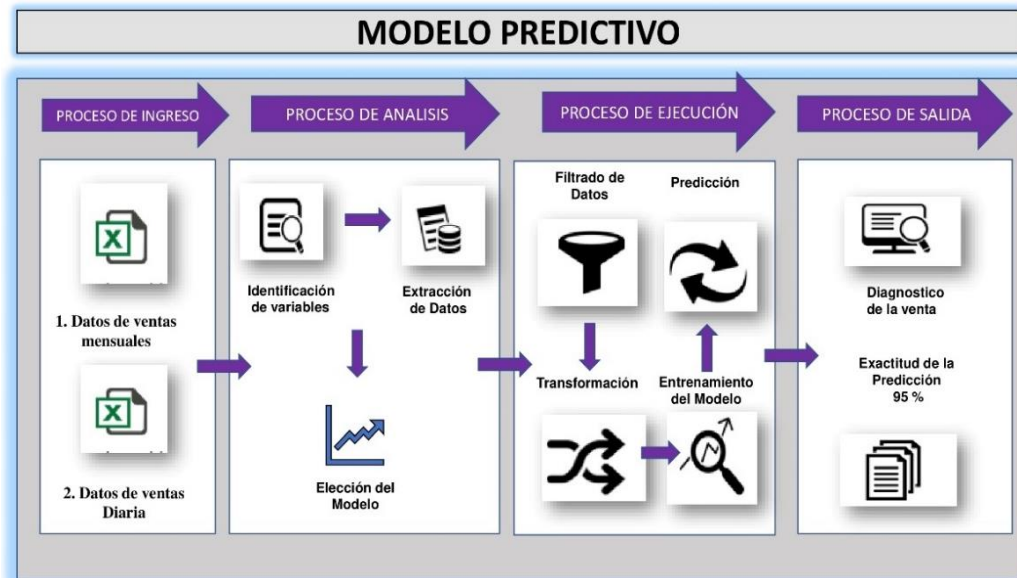
3.3.2.1. Modelo predictivo propuesto para el proceso de ventas

El análisis predictivo de ventas utiliza datos de ventas para identificar patrones en los comportamientos de clientes y clientes potenciales para realizar predicciones de ventas futuras. Por lo tanto, puede comprender mejor la salud financiera de la empresa durante los meses o años venideros. Para la presente investigación se ha propuesto un modelo predictivo considerando cuatro (04) procesos principales:

- a. Proceso de Ingreso
- b. Proceso de Análisis
- c. Proceso de Ejecución
- d. Proceso de Salida

Figura 28

Representación gráfica del modelo predictivo propuesto



Nota. Fuente, elaboración propia.

3.3.2.2. Proceso de Ingreso

El presente modelo tiene como propósito predecir las ventas futuras basándose en datos históricos. A diferencia del establecimiento de objetivos ambiciosos, los datos históricos brindan una imagen precisa y realista de la cantidad de ingresos que una empresa debería generar dentro de un período de tiempo determinado. Al relacionar datos actuales de transacciones y de interacción con el cliente con los resultados de ventas reales, se pueden llegar a predecir los ingresos futuros con un grado de precisión extremadamente alto.

Cuando se puede pronosticar con precisión cuáles serán los ingresos y detectar los riesgos con anticipación, puede usar ese conocimiento para asignar recursos y administrar la fuerza laboral de manera más eficiente. Los datos históricos de ventas también permiten comparar el desempeño de su organización con los promedios de la industria para ver si está bien encaminado. Si descubre que las métricas de ventas no se alinean con los promedios de la industria, se puede comenzar el proceso de identificar la causa raíz y hacer las correcciones necesarias.

Por tanto, en este primer proceso de ingreso se extraen los datos de las ventas del periodo de años estimado, considerando ventas diarias y mensuales, de modo que sirvan como base para tener la información a analizar para, a posteriori, llevar a cabo la predicción de ventas.

En esta fase, se hace utilizó MS-Excel para ver los atributos y datos que serían utilizados en la red neuronal a emplear.

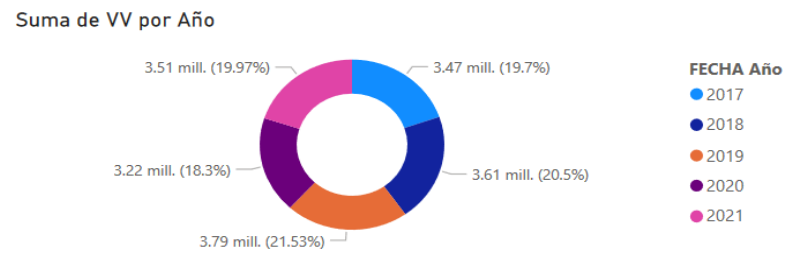
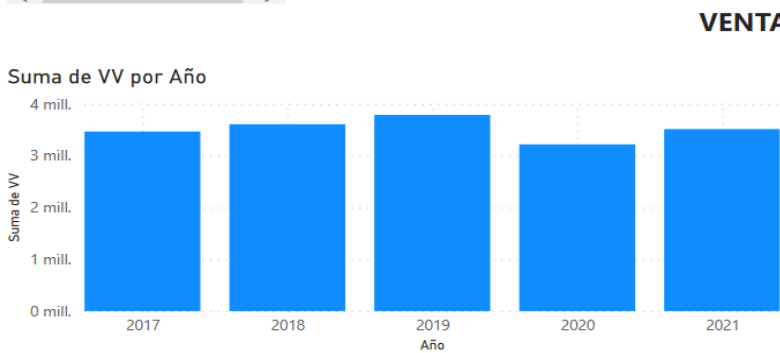
Para ello fue necesario el uso BI, haciendo el uso mediante la implementación de dashboards en el software Microsoft Power BI. Para mejorar la toma de decisiones por ello se generó los dashboards correspondientes de las ventas por año y mes desde enero 2017 hasta septiembre 2021 y realizar un análisis de los datos históricos de las ventas realizadas en la empresa textil caso de estudio. A continuación, en la Figura 29, se muestran los datos históricos de las ventas realizadas en la empresa textil caso de estudio, haciendo uso de las estrategias de BI mediante dashboards en los que se pueden visualizar las ventas por año y mes desde enero 2017 hasta septiembre 2020.

Se analizaron los datos históricos de las ventas del año 2020 y 2021 correspondientes al mes de septiembre para evaluar el comportamiento de los resultados de ventas de cada día de dicho mes para luego ser comparados con los datos estimados obtenidos posteriormente al procesar nuestra red neuronal del modelo predictivo. Tal y como se visualiza en la Figura 30.

Figura 29

Información histórica de las ventas utilizadas para el proceso de ingreso del modelo predictivo propuesto

Año	Mes	Suma de VV
2017	enero	297,410.24
2017	febrero	285,322.81
2017	marzo	282,445.06
2017	abril	231,322.76
2017	mayo	354,165.18
2017	junio	363,590.21
2017	julio	275,914.18
2017	agosto	189,710.40
2017	septiembre	258,573.15
2017	octubre	280,349.33
2017	noviembre	337,944.33
2017	diciembre	310,209.77
2018	enero	278,734.28
2018	febrero	306,732.36
2018	marzo	285,144.24
2018	abril	315,316.43
2018	mayo	359,470.09
2018	junio	302,307.20
Total		17,595,990.22



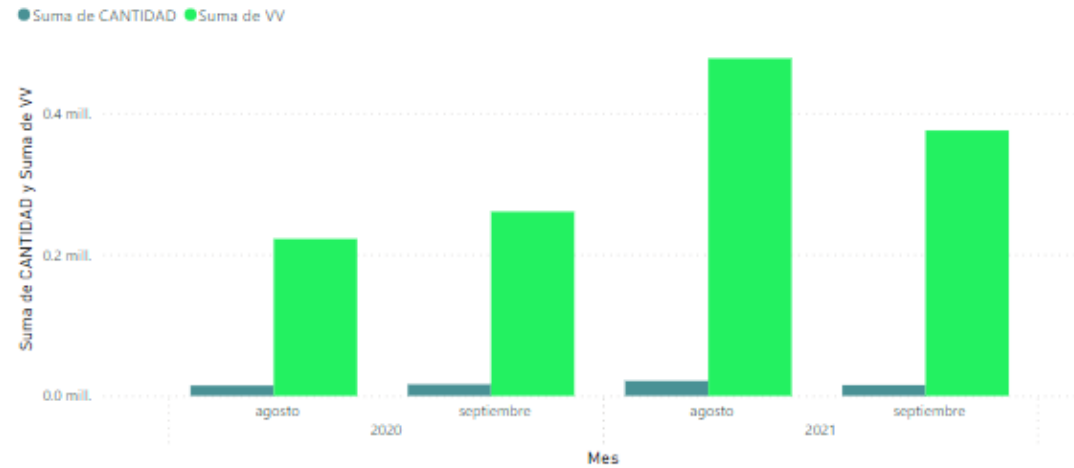
Nota. Fuente, elaboración propia.

Figura 30

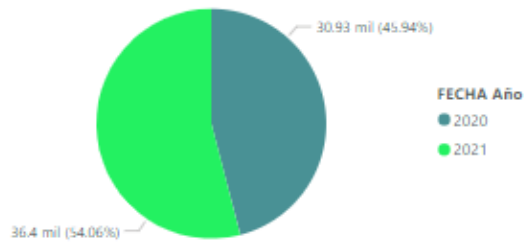
Información histórica de las ventas realizadas en los meses de agosto y septiembre 2020-2021

Año	Mes	Día	Suma de CANTIDAD
2020	agosto	1	348.53
2020	agosto	3	399.52
2020	agosto	4	631.76
2020	agosto	5	866.51
2020	agosto	6	456.26
2020	agosto	7	563.89
2020	agosto	8	405.32
2020	agosto	10	505.15
2020	agosto	11	358.59
2020	agosto	12	622.66
2020	agosto	13	440.42
2020	agosto	14	758.91
2020	agosto	15	507.69
2020	agosto	17	471.90
2020	agosto	18	325.37
Total			67,322.93

Suma de CANTIDAD y Suma de VV por Año y Mes

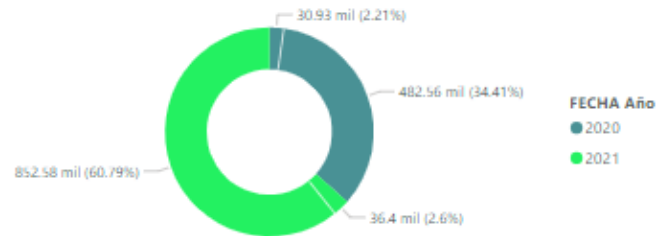


Suma de CANTIDAD por Año



VENTAS MES SETIEMBRE AÑO 2020-2021

Suma de CANTIDAD y Suma de VV por Año



3.3.2.3. Proceso de Análisis

En este proceso de análisis, lo que se busca es hacer predicciones sobre resultados futuros utilizando datos históricos combinados con modelos estadísticos, técnicas de minería de datos y/o aprendizaje automático. Las empresas emplean análisis predictivos para encontrar patrones en estos datos para identificar riesgos y oportunidades.

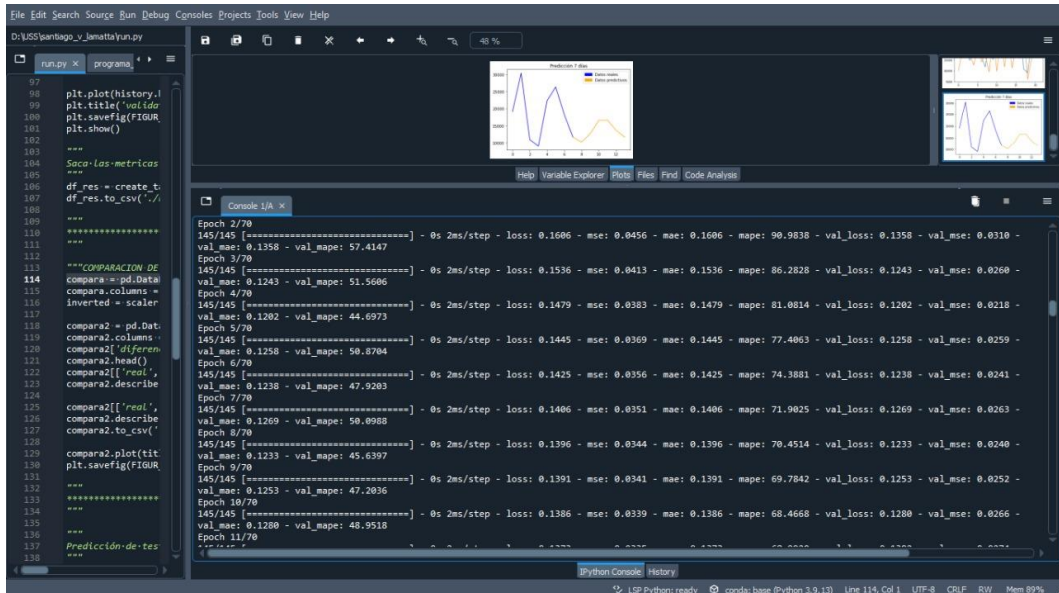
En esta fase del modelo, se busca relacionar la inteligencia de negocios y data Analytics con el propósito de lograr la predicción del pronóstico de ventas en un caso de estudio. Vale recalcar que, las empresas de hoy están se encuentran inmersas en un mar de información y datos que residen en bases de datos transaccionales, archivos de registro de equipos, imágenes, u otras fuentes de datos. Para obtener información a partir de estos datos, se utilizan modelos de aprendizaje automático con el propósito de encontrar patrones y hacer predicciones acerca de las ventas futuras. Estos incluyen regresión lineal y no lineal, redes neuronales, máquinas de vectores de soporte, árboles de decisión, entre otros. Los aprendizajes conseguidos por intermedio del análisis predictivo se pueden utilizar posteriormente dentro del análisis prescriptivo para impulsar acciones basadas en conocimientos predictivos.

Por tanto, en este proceso de análisis se pretende, primeramente, identificar las variables a utilizar, tales como fecha, día, valor de ventas, etcétera. Posteriormente se extraen los datos asociados con dichas variables y, finalmente, se hace la elección del modelo adecuado que servirá para el proceso siguiente, el cual es el proceso de ejecución del modelo.

En cuanto al diseño de datos que era necesario para lograr el almacenamiento de la data para este modelo inicialmente propuesto se hizo usanza de Business Intelligence, mediante el uso de una de sus herramientas más utilizadas, los procesos ETL. En este caso se realizó un stored procedure para recibir la información que sería alimentadora del modelo siendo útil para el uso que se propone en este caso.

Figura 31

Entrenamiento para el proceso de análisis del modelo predictivo propuesto



Nota. Fuente, elaboración propia.

3.3.2.4. Proceso de Ejecución

Este proceso de ejecución tiene que ver con llevar a la praxis el modelo predictivo. En este proceso se siguen una serie de actividades descritas a continuación:

a. Actividad de filtrado de datos

En esta actividad se verifica si existen valores nulos, erróneos y/o cualquier otro tipo de datos inconsistentes, de modo que luego puedan ser depurados quedando solamente aquellos datos que sí sean adecuados para poder transformarlos posteriormente.

b. Actividad de transformación

En esta actividad, los datos a utilizar que previamente fueron filtrados, son transformados mediante un script de lenguaje de programación seleccionado según criterios, para posteriormente poder ser empleados con dicho algoritmo el cual le brinde mejores resultados en la predicción de ventas.

c. Actividad del modelo elección del modelo

En esta actividad, primeramente, se seleccionaron la cantidad de datos a entrenar para posteriormente, utilizar el modelo de aprendizaje automático a utilizar.

d. Actividad de predicción

En esta actividad, una vez aplicado el algoritmo seleccionado previamente, el cual mostró resultados precisos en cuanto a la predicción de procesos de ventas, este es utilizado para hacer las predicciones en torno al proceso de ventas de un caso de estudio seleccionado.

3.3.2.5. Proceso de Salida

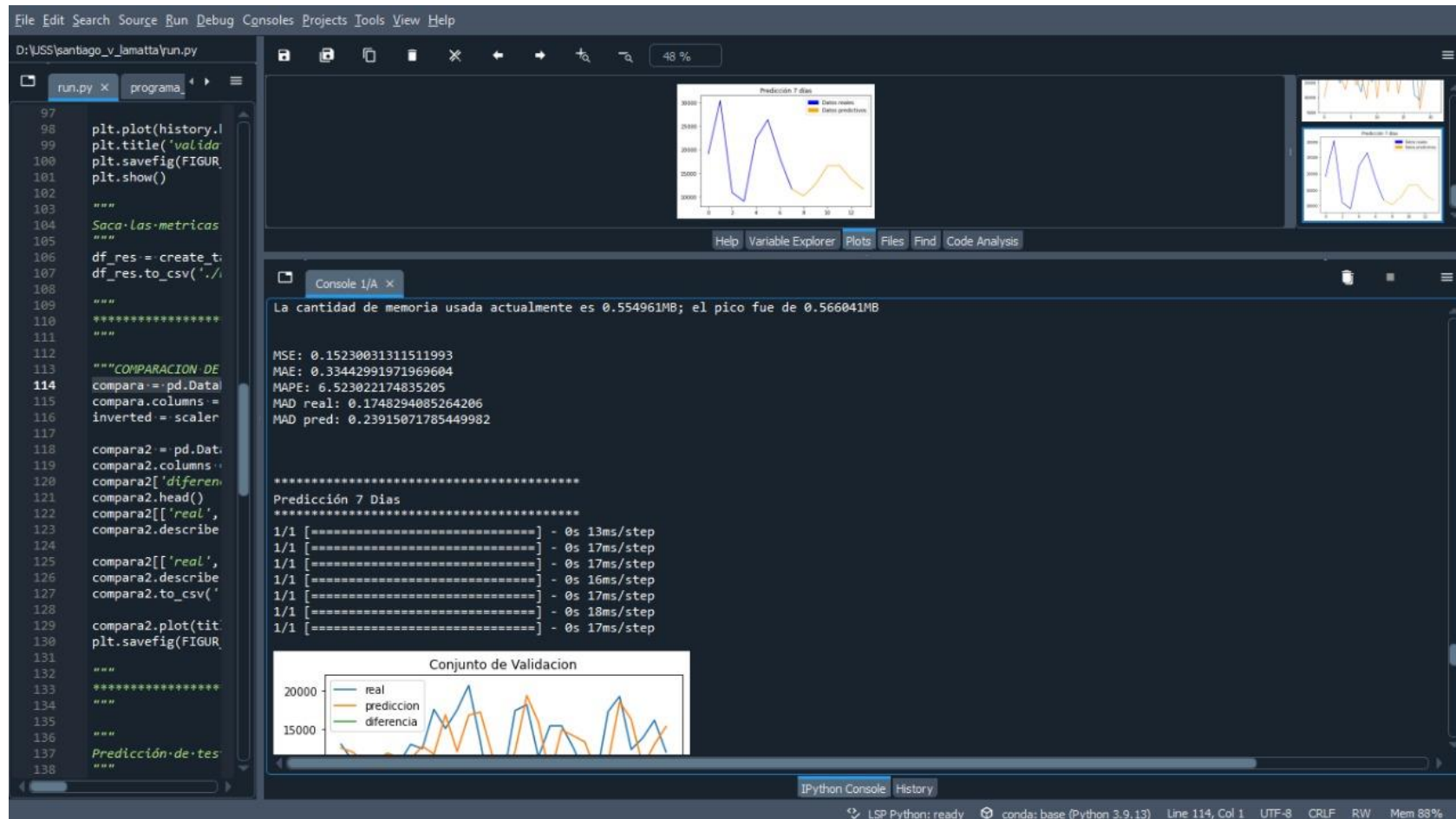
Finalmente, en este proceso de salida, lo que se lleva a cabo es tener los resultados del diagnóstico de predicción favorables de las actividades desarrolladas previamente, de tal manera que se puedan documentar los pronósticos del proceso de ventas que utilice la empresa para predecir sus ventas futuras.

Al realizar el análisis del comportamiento de los datos y evaluación de los resultados de los años 2017, 2018, 2019, 2020, 2021. Para la toma de decisiones se efectuó un nuevo análisis del comportamiento del mes de

setiembre del año 2020 y 2021 como se muestra en la Figura 33, donde se visualizan los valores de correspondientes de los días 16 al 30 del mes de setiembre del año 2020 y de los días 16 al 23 del año 2021. Con dicho análisis se necesitaba realizar la predicción de las ventas de los 7 días a partir del día 24 al 30 de setiembre del año 2021, para la toma de decisiones correspondiente.

Figura 32

Predicción para el proceso de análisis del modelo predictivo propuesto



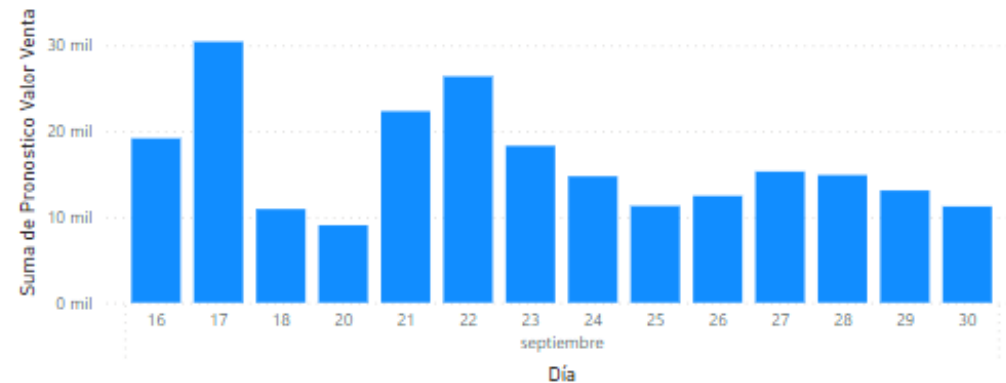
Nota. Fuente, elaboración propia.

Figura 33

Comparación de datos históricos vs datos predictivos

Año	Mes	Día	Suma de Pronostico Valor Venta
2021	septiembre	22	26,389.45
2021	septiembre	21	22,331.15
2021	septiembre	16	19,164.23
2021	septiembre	23	18,292.25
2021	septiembre	27	15,296.98
2021	septiembre	28	14,878.19
2021	septiembre	24	14,739.48
2021	septiembre	29	13,088.73
2021	septiembre	26	12,456.31
2021	septiembre	25	11,279.25
2021	septiembre	30	11,221.94
2021	septiembre	18	10,916.04
2021	septiembre	20	9,026.93
Total			229,523.41

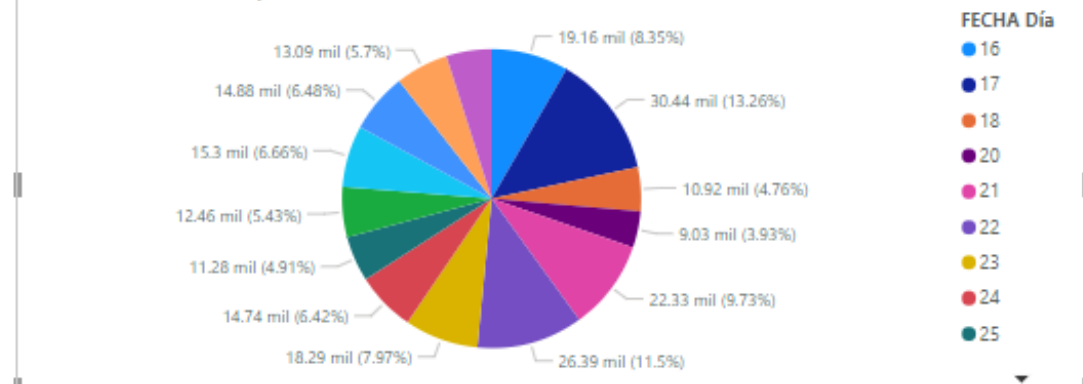
Suma de Pronostico Valor Venta por Mes y Día



**PREDICION DE VENTAS DE 7 DIAS
MES SETIEMBRE 2021**

Año	Mes	Día	Suma de PREDICION	VALOR VENTA REAL
2021	septiembre	16	19,164.23	19164.23
2021	septiembre	17	30,442.47	30442.47
2021	septiembre	18	10,916.04	10916.04
2021	septiembre	20	9,026.93	9026.93
2021	septiembre	21	22,331.15	22331.15
2021	septiembre	22	26,389.45	26389.45
2021	septiembre	23	18,292.25	18292.25
2021	septiembre	24	14,739.48	-
2021	septiembre	25	11,279.25	-
2021	septiembre	26	12,456.31	-
2021	septiembre	27	15,296.98	-
2021	septiembre	28	14,878.19	-
2021	septiembre	29	13,088.73	-
2021	septiembre	30	11,221.94	-
Total			229,523.41	

Suma de PREDICION por Día



3.3.3. Implementación del modelo predictivo

3.3.3.1. Arquitectura del modelo

El propósito de este informe fue, primeramente, la construcción del modelo automatizado aplicando redes neuronales, para efectuar el pronóstico de ventas de la empresa Centro Textil de la Matta.

Teniendo en cuenta los datos de las ventas diarias (del último lustro), para pronosticar las ventas de cada día tomando en consideración los datos históricos de los 7 días anteriores. Para poder predecir las ventas diarias, se realizó un análisis de regresión de series de tiempo haciendo usanza de información de desplazamiento histórica combinada con una red neuronal.

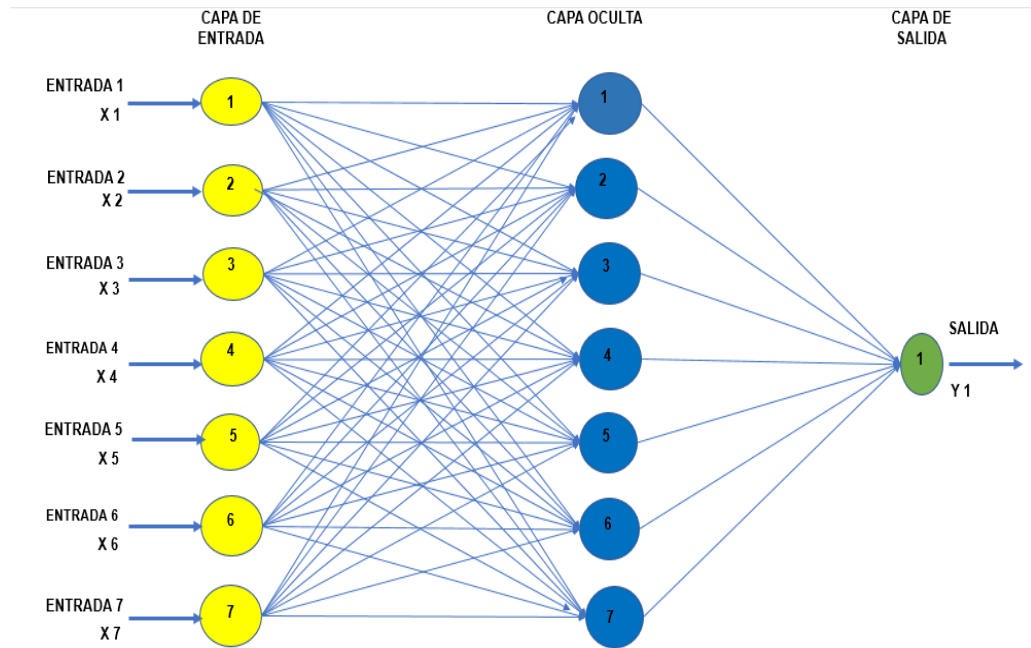
En efecto se ejemplifica la arquitectura de la red neuronal que se aplicó para el proceso del pronóstico de las ventas.

1. Capas de entrada 7
2. 1 capa oculta con 7 neuronas (este valor podría modificarse)
3. La salida será de 1 sola neurona.
4. Como función de activación se utilizó tangente hiperbólica para lo cual se utiliza valores entre -1 y 1
5. Se utilizará como optimizador ADAM y métrica de pérdida(los) Mean Absolute Error.
6. La predicción será con un valor continuo y no discreto para poder computar el Acuracy utilizando Mean Squared error donde se realizará el entrenamiento de 70 épocas.

A continuación, se muestra el esquema de la Red Neuronal para la solución del pronóstico de las ventas:

Figura 34

Esquema de la Red Neuronal para la solución del pronóstico de las ventas



Nota. Fuente, elaboración propia.

En la **Figura 33** se muestra la implementación en el lenguaje Python de la arquitectura de red neuronal:

Figura 35

Script de la arquitectura de la red neuronal en Python

```
Def crear_modeloff():  
    Model = Sequential()  
    Model.add(Dense(PASOS, input_shape=(1,PASOS),activation='tanh'))  
    Model.add(Flatten())  
    Model.add(Dense(1, activation='tanh'))
```

Nota. Fuente, elaboración propia.

3.3.3.2. Lectura de la codificación del proceso del modelo predictivo

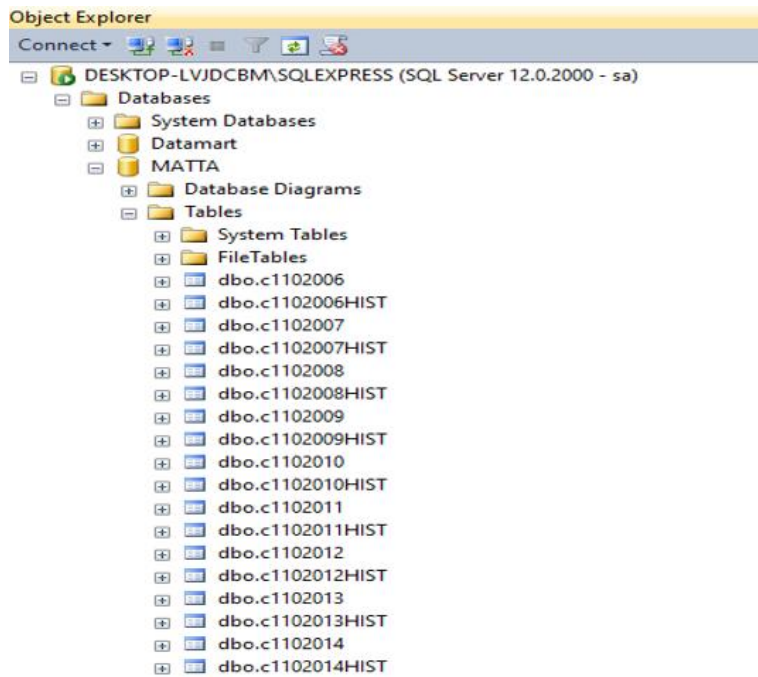
A continuación, se detalla el proceso de la implementación del modelo predictivo utilizando como base de datos SQL Server, lenguaje de programación visual FoxPro 2.6 y Python para el funcionamiento y proceso del pronóstico de ventas.

3.3.3.3. Adquisición de los datos

Para la adquisición de los datos se obtuvo DB SQL de nombre MATTA que se encuentra alojada en el servidor de base de datos. Como se muestra en la figura siguiente.

Figura 36

Modelo de base datos DB SQL Matta



Nota. Fuente, elaboración propia.

Obteniendo el conjunto de dato pertenecientes a las ventas de los años 2017, 2018, 2019, 2020 y 2021, es decir, de los últimos cinco (05) años. Para dicha actividad, se necesitó de un script realizado mediante lenguaje de programación Visual Fox-Pro para poder hacer la transacción de la BD SQL realizando la extracción de los datos del servidor hacia una hoja de cálculo MS-Excel, el cual se muestra en la Figura 37, para luego ser procesada en Python. A continuación, se muestra el Script para transacción BD SQL Server a Excel:

Figura 37

Stored procedure para transacción BD SQL Server a Excel

```
FOR j=2017 TO 2021

    g_ano=PADL(j,4,'0')

    l1="insert into #rvta select c210.fecha, YEAR(c210.fecha) anio, MONTH(c210.fecha) mes,
SUM(c215.soles*case when c215.tipo='2' then 1 else -1 end) soles "

    l2="from c210"+g_ano+" c210,c215"+g_ano+" c215 where c210.llave=c215.llave and
LEFT(c215.cuenta,2)='70' and c210.libro='14' group by c210.fecha "
```

```

=ej(l1+l2,"
ENDFOR
nlin=1
l1="select anio,mes,000 numero,SUM(soles) soles from #rvta group by anio,mes order by
anio,numero "
=ej(l1,'vta')
GO top
SCAN WHILE !EOF('vta')
    rEPLACE vta.numero WITH nlin
    nlin=nlin+1
ENDSCAN
SELECT vta

```

Nota. Fuente, elaboración propia.

En este caso, se han obtenido la totalidad de los datos referentes a las ventas históricas de los años mencionados anteriormente, por lo que se han podido distinguir diez (10) variables que son las siguientes:

Tabla 17
Variables existentes en las ventas históricas

N°	Variable	Descripción
1	ANIO	Indica el año de la venta
2	MES	Indica el mes de la venta
3	DOCU	Indica el tipo de documento de la venta, ya sea boleta de venta o factura
4	FECHA	Indica la fecha de venta
5	CTACTE	Indica el código interno asignado a un cliente
6	RAZON	Indica la razón social del cliente

7	CODIGO	Indica el código del producto expendido
8	DESCRIP	Indica la descripción del producto
9	VV	Indica el valor de venta del producto expresado en soles
10	CANT	Indica la cantidad de productos expendidos

Nota. Fuente, elaboración propia.

3.3.3.4. Selección de variables

Para la selección de variables, se tuvieron que seleccionar aquellas que nos permitan determinar los resultados para poder considerar la predicción de ventas. Para ello se desarrolló un ranking considerando ciertas características que debían cumplir dichas variables. Dichas características se muestran a continuación:

Tabla 18

Ranking de valores para selección de variables a utilizar

N°	Variable	Medible	Predictivo	Estadístico	Necesario	TOTAL
1	ANIO	x	x	x		3
2	MES	x	x	x		3
3	DOCU					0
4	FECHA	x	x	x	x	4
5	CTACTE					0
6	RAZON					0
7	CODIGO					0
8	DESCRIP					0
9	VV	x	x	x	x	4
10	CANT	x	x	x		3

Nota. Fuente, elaboración propia.

Posteriormente, se seleccionaron, de entre las diez (10) variables iniciales, tan solamente a dos (02) de las mismas: las variables “FECHA” y “VV”. Dichas variables sirvieron para hacer la predicción de las ventas de la empresa caso de estudio. Por tanto, el conjunto de datos posterior a la limpieza y selección de variables quedó constituido por 4383 datos ya depurados.

Figura 38

Datos obtenidos posteriores a la limpieza de datos

	A	C	D	E
1	FECHA	VV		
2	2/01/2017	156.66		
3	3/01/2017	4781.79		
4	4/01/2017	7563.97		
5	5/01/2017	11155.98		
6	6/01/2017	8664.01		
7	7/01/2017	12925.96		
8	9/01/2017	4337.70		
9	10/01/2017	13804.76		
10	11/01/2017	11185.96		
11	12/01/2017	11195.48		
12	13/01/2017	16397.61		
13	14/01/2017	4246.08	.	
14	16/01/2017	8238.41		
15	17/01/2017	10452.31		
16	18/01/2017	11278.46		
17	19/01/2017	6118.21		
18	20/01/2017	30602.32		
19	21/01/2017	6079.67		
20	23/01/2017	20326.68		
21	24/01/2017	18359.18		
22	25/01/2017	16406.53		
23	26/01/2017	11323.71		
24	27/01/2017	17864.56		
25	28/01/2017	1414.71		
26	29/01/2017	11060.00		

Nota. Fuente, elaboración propia.

3.3.3.5. Proceso de inicio en python

Se utilizó el lenguaje de programación Python utilizando como plataforma Spyder Al inicio de la ejecutar se realizó el proceso de importar las librerías pandas, numpy y matplotlib.

Figura 39

Proceso de Inicio en Python (importación de librerías a utilizar)

```
"""Cargamos nuestras librerias necesarias de uso
para nuestro modelo
"""
import pandas as pd
import numpy as np
import psutil
import tracemalloc
import matplotlib.pyplot as plt
```

Nota. Fuente, elaboración propia.

3.3.3.6. Cargamos Nuestro DATASET

Se realiza la carga de datos de las ventas diarias por día de los años 2017-2018-2019-2020-2021.

Figura 40

Script cargar de datos del dataset en python

```
"""Cargamos nuestro Dataset
Prepara el dataset raw convirtiendolo a csv para poder manejarlo
con
mayor facilidad y extrae las columnas necesarias.
"""
xlsx_to_csv(INPUT_FILE, sheet_name='venta_diaria',
cols=['FECHA', 'VV'], output=INTER_FILE)
split_train_test(INTER_FILE,
train_range=["2017-01-02", "2020-12-31"],
test_range=["2021-01-02", "2021-09-23"],
output_path=INTER_PATH)
```

Nota. Fuente, elaboración propia.

Figura 41

Proceso de cálculo de valores de venta por día acumulado

```
def build_feature(input, pasos, output):
df = pd.read_csv(input).set_index("FECHA")
df.index = pd.to_datetime(df.index)
# Creamos un indice y asignamos como indice al campo Fecha, por que
trabajaremos con serie de tiempo
df = df.resample('D').mean()
df = df.rolling(2).mean()
# Verificar valores nulos
df.loc[df['VV'].isnull()]
# Detectar valores extraños
df[df['VV'].isin([np.nan, np.inf, -np.inf])]
# Reemplazar valores NaN
df['VV'] = df['VV'].fillna(0)
df = df.drop(df[df['VV']==0].index)
# load dataset
values = df.values
# ensure all data is float
values = values.astype('float32')
# normalize features
scaler = MinMaxScaler(feature_range=(-1, 1))
values = values.reshape(-1, 1) # esto lo hacemos porque tenemos 1 sola
dimension
scaled = scaler.fit_transform(values)
# frame as supervised learning
reframed = series_to_supervised(scaled, pasos, 1)
reframed.head().to_csv(REPORT_DATA_PATH + "feature_head.csv")
reframed.describe().to_csv(REPORT_DATA_PATH + "feature_describe.csv")
reframed.to_csv(output, index=False)
```

Nota. Fuente, elaboración propia.

Tabla 19

Datos de los años 2017-2018-2019-2020-2021 de valores para selección de variables a utilizar

AÑO 2017		AÑO 2018		AÑO 2019		AÑO 2020		AÑO 2021	
FECHA	VV (VALOR VENTA)	FECHA	VV (VALOR VENTA)	FECHA	VV (VALOR VENTA)	FECHA	VV (VALOR VENTA)	FECHA	VV (VALOR VENTA)
02/01/2017	156.66	02/01/2018	2,015.38	02/01/2019	8,959.93	06/01/2020	8,475.40	02/01/2021	2.53
03/01/2017	4,781.79	03/01/2018	4,992.91	03/01/2019	7,840.06	07/01/2020	32,487.62	04/01/2021	5,639.82
04/01/2017	7,563.97	04/01/2018	4,907.29	04/01/2019	6,542.96	08/01/2020	35,098.90	05/01/2021	9,532.55
05/01/2017	11,155.98	05/01/2018	8,901.84	05/01/2019	2,378.12	09/01/2020	17,665.78	06/01/2021	12,642.46
06/01/2017	8,664.01	06/01/2018	2,277.03	07/01/2019	10,883.07	10/01/2020	31,472.86	07/01/2021	17,690.63
07/01/2017	12,925.96	08/01/2018	10,798.77	08/01/2019	11,151.19	11/01/2020	19,103.18	08/01/2021	29,231.02
09/01/2017	4,337.70	09/01/2018	12,893.18	09/01/2019	21,736.34	13/01/2020	25,854.14	09/01/2021	10,634.95
10/01/2017	13,804.76	10/01/2018	13,026.20	10/01/2019	16,511.41	14/01/2020	39,476.62	11/01/2021	8,359.94
11/01/2017	11,185.96	11/01/2018	7,114.91	11/01/2019	12,108.93	15/01/2020	31,156.78	12/01/2021	14,717.28
12/01/2017	11,195.48	12/01/2018	13,962.04	12/01/2019	13,584.33	16/01/2020	31,940.84	13/01/2021	21,180.59
13/01/2017	16,397.61	13/01/2018	6,203.99	14/01/2019	10,342.77	17/01/2020	33,441.10	14/01/2021	35,073.01
14/01/2017	4,246.08	15/01/2018	14,294.35	15/01/2019	9,148.89	18/01/2020	15,872.86	15/01/2021	11,684.38
16/01/2017	8,238.41	16/01/2018	8,439.68	16/01/2019	12,628.22	20/01/2020	38,177.22	16/01/2021	8,057.13
17/01/2017	10,452.31	17/01/2018	14,888.44	17/01/2019	13,978.97	21/01/2020	38,317.14	18/01/2021	10,361.44

18/01/2017	11,278.46	18/01/2018	14,744.77	18/01/2019	14,821.61	22/01/2020	36,978.04	19/01/2021	12,611.59
19/01/2017	6,118.21	19/01/2018	12,960.03	19/01/2019	4,296.27	23/01/2020	36,687.52	20/01/2021	12,154.37
20/01/2017	30,602.32	20/01/2018	9,774.32	20/01/2019	1,322.89	24/01/2020	53,530.52	21/01/2021	12,637.12
21/01/2017	6,079.67	22/01/2018	8,389.20	21/01/2019	10,717.05	25/01/2020	16,598.50	22/01/2021	13,343.24
23/01/2017	20,326.68	23/01/2018	18,047.07	22/01/2019	13,195.41	27/01/2020	42,395.76	23/01/2021	8,591.32
24/01/2017	18,359.18	24/01/2018	11,991.58	23/01/2019	9,581.21	28/01/2020	26,326.34	24/01/2021	1,702.14
25/01/2017	16,406.53	25/01/2018	10,247.46	24/01/2019	14,963.51	29/01/2020	28,970.96	25/01/2021	21,788.51
26/01/2017	11,323.71	26/01/2018	12,902.34	25/01/2019	9,858.86	30/01/2020	43,817.62	26/01/2021	13,639.88
27/01/2017	17,864.56	27/01/2018	10,943.76	26/01/2019	4,511.73	31/01/2020	43,883.62	27/01/2021	14,824.41
28/01/2017	1,414.71	29/01/2018	11,134.01	28/01/2019	10,372.32	01/02/2020	7,173.85	28/01/2021	17,354.76
30/01/2017	11,060.99	30/01/2018	15,230.39	29/01/2019	8,533.50	03/02/2020	7,885.16	29/01/2021	16,448.40
31/01/2017	21,468.54	31/01/2018	17,653.34	30/01/2019	16,096.44	04/02/2020	14,974.76	30/01/2021	10,846.24
01/02/2017	14,703.29	01/02/2018	25,088.92	31/01/2019	10,797.58	05/02/2020	10,382.76	01/02/2021	10,126.33
02/02/2017	14,590.48	02/02/2018	20,737.10	01/02/2019	11,162.08	06/02/2020	17,258.76	02/02/2021	22,091.88
03/02/2017	14,013.24	03/02/2018	10,640.49	02/02/2019	3,425.56	07/02/2020	15,361.83	03/02/2021	10,720.18
04/02/2017	6,540.58	04/02/2018	32.31	04/02/2019	3,495.45	08/02/2020	3,880.86	04/02/2021	10,183.93
06/02/2017	10,700.90	05/02/2018	14,301.12	05/02/2019	8,751.98	09/02/2020	0.00	05/02/2021	6,755.16
07/02/2017	12,676.44	06/02/2018	11,926.60	06/02/2019	8,553.08	10/02/2020	12,891.61	06/02/2021	9,179.60
08/02/2017	16,153.21	07/02/2018	16,107.78	07/02/2019	8,756.83	11/02/2020	13,735.30	08/02/2021	10,392.61

09/02/2017	13,886.61	08/02/2018	8,945.92	08/02/2019	15,595.83	12/02/2020	23,337.44	09/02/2021	12,905.41
10/02/2017	15,885.61	09/02/2018	22,779.74	09/02/2019	3,060.05	13/02/2020	6,379.15	10/02/2021	11,805.73
11/02/2017	6,867.60	10/02/2018	1,885.47	11/02/2019	10,494.49	14/02/2020	11,121.05	11/02/2021	16,023.36
13/02/2017	9,846.87	12/02/2018	11,472.60	12/02/2019	14,482.34	15/02/2020	5,147.17	12/02/2021	10,852.93
14/02/2017	6,805.35	13/02/2018	11,206.95	13/02/2019	6,006.85	17/02/2020	11,744.11	13/02/2021	5,080.77
15/02/2017	9,123.42	14/02/2018	9,921.23	14/02/2019	7,401.28	18/02/2020	16,191.61	15/02/2021	14,011.90
16/02/2017	8,468.46	15/02/2018	10,629.65	15/02/2019	12,220.43	19/02/2020	24,279.32	16/02/2021	9,808.22
17/02/2017	12,341.55	16/02/2018	15,578.02	16/02/2019	6,318.13	20/02/2020	16,716.97	17/02/2021	19,188.66
18/02/2017	3,394.73	17/02/2018	3,729.80	17/02/2019	34.30	21/02/2020	16,607.60	18/02/2021	25,349.78
20/02/2017	8,242.00	19/02/2018	7,344.30	18/02/2019	14,158.32	22/02/2020	7,505.96	19/02/2021	13,736.49
21/02/2017	13,589.01	20/02/2018	11,133.38	19/02/2019	11,190.97	24/02/2020	13,352.64	20/02/2021	16,406.37
22/02/2017	17,651.96	21/02/2018	10,432.20	20/02/2019	9,342.82	25/02/2020	23,600.18	21/02/2021	191.40
23/02/2017	13,513.71	22/02/2018	14,248.69	21/02/2019	6,126.81	26/02/2020	20,792.38	22/02/2021	7,230.39
24/02/2017	13,083.14	23/02/2018	15,106.26	22/02/2019	7,815.40	27/02/2020	23,603.90	23/02/2021	15,521.82
25/02/2017	6,863.76	24/02/2018	5,748.48	23/02/2019	10,433.89	28/02/2020	16,269.36	24/02/2021	11,006.09
27/02/2017	19,678.13	26/02/2018	12,569.22	25/02/2019	6,763.48	29/02/2020	11,833.15	25/02/2021	10,999.83
28/02/2017	16,702.76	27/02/2018	11,232.46	26/02/2019	14,168.95	02/03/2020	8,123.02	26/02/2021	9,556.41
01/03/2017	13,107.36	28/02/2018	23,933.67	27/02/2019	6,940.53	03/03/2020	21,058.06	27/02/2021	9,960.45
02/03/2017	13,350.57	01/03/2018	15,086.12	28/02/2019	9,017.21	04/03/2020	10,286.27	28/02/2021	3,967.27

03/03/2017	11,310.94	02/03/2018	11,942.23	01/03/2019	6,974.54	05/03/2020	17,809.59	01/03/2021	27,368.09
04/03/2017	2,111.94	03/03/2018	4,853.61	02/03/2019	3,521.23	06/03/2020	15,493.21	02/03/2021	361.34
06/03/2017	19,280.86	05/03/2018	12,401.70	04/03/2019	7,429.86	07/03/2020	8,004.43	03/03/2021	22,780.30
07/03/2017	14,487.08	06/03/2018	10,285.57	05/03/2019	15,053.69	09/03/2020	12,158.72	04/03/2021	12,391.28
08/03/2017	8,337.44	07/03/2018	8,944.13	06/03/2019	7,192.15	10/03/2020	18,400.59	05/03/2021	20,064.71
09/03/2017	7,101.44	08/03/2018	11,509.60	07/03/2019	7,513.50	11/03/2020	9,973.66	06/03/2021	17,841.69
10/03/2017	10,519.34	09/03/2018	11,470.15	08/03/2019	5,406.17	12/03/2020	7,877.82	08/03/2021	12,612.79
11/03/2017	3,181.62	10/03/2018	5,011.44	09/03/2019	3,348.86	13/03/2020	12,485.97	09/03/2021	15,229.67
13/03/2017	10,837.03	12/03/2018	12,090.26	11/03/2019	6,279.57	22/03/2020	1,625.89	10/03/2021	28,083.41
14/03/2017	7,614.89	13/03/2018	10,972.00	12/03/2019	28,983.92	25/03/2020	8.49	11/03/2021	9,754.39
15/03/2017	18,979.14	14/03/2018	7,043.51	13/03/2019	7,997.86	30/03/2020	1.43	12/03/2021	9,946.30
16/03/2017	11,692.33	15/03/2018	12,277.90	14/03/2019	7,188.45	15/04/2020	0.00	13/03/2021	17,325.04
17/03/2017	8,009.47	16/03/2018	13,908.07	15/03/2019	12,445.02	16/04/2020	6,911.99	15/03/2021	5,622.31
18/03/2017	2,186.46	17/03/2018	9,098.91	16/03/2019	3,013.89	17/04/2020	3,417.29	16/03/2021	20,717.67
20/03/2017	9,432.56	19/03/2018	12,686.25	18/03/2019	9,098.91	20/04/2020	49.01	17/03/2021	7,852.70
21/03/2017	19,874.91	20/03/2018	6,574.03	19/03/2019	13,710.73	21/04/2020	552.67	18/03/2021	8,453.91
22/03/2017	4,669.74	21/03/2018	11,054.35	20/03/2019	5,640.18	22/04/2020	2,672.11	19/03/2021	15,780.83
23/03/2017	8,295.53	22/03/2018	9,966.69	21/03/2019	18,300.81	23/04/2020	239.16	20/03/2021	10,640.39
24/03/2017	7,870.18	23/03/2018	18,098.28	22/03/2019	11,085.58	24/04/2020	122.86	21/03/2021	1,219.36

25/03/2017	5,146.48	24/03/2018	9,438.42	23/03/2019	6,134.12	25/04/2020	742.36	22/03/2021	10,991.37
27/03/2017	9,116.00	26/03/2018	15,396.14	24/03/2019	423.49	27/04/2020	251.15	23/03/2021	15,530.11
28/03/2017	11,851.90	27/03/2018	16,532.13	25/03/2019	10,836.66	28/04/2020	1,313.64	24/03/2021	23,964.71
29/03/2017	16,350.36	28/03/2018	18,999.13	26/03/2019	13,399.49	29/04/2020	866.54	25/03/2021	26,182.03
30/03/2017	10,686.39	29/03/2018	3,591.40	27/03/2019	12,908.30	30/04/2020	1,875.73	26/03/2021	12,381.83
31/03/2017	17,043.10	30/03/2018	6.68	28/03/2019	11,653.43	04/05/2020	2,576.79	27/03/2021	7,639.22
01/04/2017	4,200.05	31/03/2018	5,905.54	29/03/2019	9,480.71	05/05/2020	2,605.21	29/03/2021	8,106.86
03/04/2017	9,881.33	02/04/2018	12,738.78	30/03/2019	3,361.56	06/05/2020	727.73	30/03/2021	22,366.00
04/04/2017	13,373.05	03/04/2018	18,505.31	01/04/2019	9,704.88	07/05/2020	1,419.28	31/03/2021	10,574.14
05/04/2017	13,727.85	04/04/2018	18,660.92	02/04/2019	14,631.57	08/05/2020	1,510.83	05/04/2021	17,226.36
06/04/2017	8,828.49	05/04/2018	9,694.46	03/04/2019	10,628.09	11/05/2020	2,442.01	06/04/2021	10,971.77
07/04/2017	7,752.73	06/04/2018	12,274.20	04/04/2019	11,135.09	12/05/2020	3,407.09	07/04/2021	21,818.56
08/04/2017	4,137.48	07/04/2018	9,669.84	05/04/2019	6,922.66	13/05/2020	2,823.05	08/04/2021	33,755.40
10/04/2017	8,148.59	09/04/2018	5,320.35	06/04/2019	9,380.99	14/05/2020	4,392.61	09/04/2021	15,393.61
11/04/2017	4,404.28	10/04/2018	7,774.33	08/04/2019	9,475.33	15/05/2020	8,709.09	10/04/2021	20,543.85
12/04/2017	10,500.22	11/04/2018	22,312.54	09/04/2019	6,761.08	16/05/2020	1,823.52	12/04/2021	5,342.62
13/04/2017	418.92	12/04/2018	6,691.59	10/04/2019	37,219.72	17/05/2020	625.04	13/04/2021	11,617.71
14/04/2017	78.28	13/04/2018	16,956.06	11/04/2019	7,636.65	18/05/2020	3,834.69	14/04/2021	14,671.94
15/04/2017	17.13	14/04/2018	3,710.62	12/04/2019	7,118.11	19/05/2020	9,487.46	15/04/2021	17,584.65

17/04/2017	10,550.38	16/04/2018	11,798.03	13/04/2019	4,685.51	20/05/2020	9,253.74	16/04/2021	13,780.33
18/04/2017	5,129.17	17/04/2018	13,610.29	14/04/2019	1,173.18	21/05/2020	5,317.30	17/04/2021	7,339.38
19/04/2017	8,307.55	18/04/2018	14,086.07	15/04/2019	12,635.12	22/05/2020	11,555.68	19/04/2021	17,022.17
20/04/2017	13,795.32	19/04/2018	12,087.81	16/04/2019	13,033.69	23/05/2020	7,386.79	20/04/2021	11,776.92
21/04/2017	12,199.73	20/04/2018	12,422.16	17/04/2019	26,239.28	25/05/2020	9,620.25	21/04/2021	8,179.75
22/04/2017	6,223.57	21/04/2018	3,850.28	18/04/2019	4,626.31	26/05/2020	5,854.41	22/04/2021	12,175.22
24/04/2017	14,988.22	22/04/2018	246.21	20/04/2019	979.56	27/05/2020	10,323.53	23/04/2021	27,401.57
25/04/2017	13,680.47	23/04/2018	14,845.60	21/04/2019	0.00	28/05/2020	15,554.04	24/04/2021	15,011.05
26/04/2017	15,260.92	24/04/2018	12,089.54	22/04/2019	6,851.69	29/05/2020	4,153.05	26/04/2021	8,777.12
27/04/2017	14,389.02	25/04/2018	20,700.24	23/04/2019	10,742.62	30/05/2020	12,584.62	27/04/2021	9,947.30
28/04/2017	21,480.35	26/04/2018	12,063.20	24/04/2019	18,696.89	01/06/2020	12,498.08	28/04/2021	32,390.54
29/04/2017	9,849.66	27/04/2018	18,881.37	25/04/2019	22,523.03	02/06/2020	8,022.42	29/04/2021	24,205.54
01/05/2017	1,351.13	28/04/2018	4,053.44	26/04/2019	15,463.97	03/06/2020	7,677.86	30/04/2021	25,048.34
02/05/2017	14,083.41	29/04/2018	207.25	27/04/2019	12,588.92	04/06/2020	7,347.55	03/05/2021	7,517.37
03/05/2017	13,915.48	30/04/2018	20,065.94	28/04/2019	927.94	05/06/2020	9,267.36	04/05/2021	11,059.04
04/05/2017	13,471.07	02/05/2018	17,749.59	29/04/2019	23,318.75	06/06/2020	5,240.51	05/05/2021	9,659.25
05/05/2017	7,675.45	03/05/2018	18,719.83	30/04/2019	25,866.18	08/06/2020	12,322.08	06/05/2021	12,674.78
06/05/2017	5,818.01	04/05/2018	14,318.26	02/05/2019	9,264.63	09/06/2020	11,553.51	07/05/2021	11,093.82
08/05/2017	9,458.98	05/05/2018	7,360.53	03/05/2019	16,902.80	10/06/2020	6,897.43	08/05/2021	9,982.20

09/05/2017	19,392.03	06/05/2018	90.28	04/05/2019	4,069.27	11/06/2020	9,003.88	10/05/2021	6,103.90
10/05/2017	22,234.78	07/05/2018	8,685.10	05/05/2019	1,368.66	12/06/2020	5,190.05	11/05/2021	9,379.87
11/05/2017	24,099.92	08/05/2018	20,330.19	06/05/2019	14,949.57	13/06/2020	5,796.13	12/05/2021	10,929.77
12/05/2017	7,842.39	09/05/2018	13,628.71	07/05/2019	8,507.23	15/06/2020	10,625.22	13/05/2021	12,067.29
13/05/2017	12,236.70	10/05/2018	13,044.76	08/05/2019	17,984.04	16/06/2020	9,411.51	14/05/2021	10,297.60
15/05/2017	20,382.67	11/05/2018	7,635.83	09/05/2019	18,138.03	17/06/2020	8,058.82	15/05/2021	8,050.71
16/05/2017	11,552.32	12/05/2018	4,641.85	10/05/2019	14,782.96	18/06/2020	8,812.95	17/05/2021	9,475.40
17/05/2017	13,962.99	14/05/2018	7,143.56	11/05/2019	6,396.41	19/06/2020	5,505.64	18/05/2021	16,858.47
18/05/2017	14,634.85	15/05/2018	14,708.83	13/05/2019	10,043.03	20/06/2020	4,533.95	19/05/2021	12,013.84
19/05/2017	15,797.54	16/05/2018	14,983.04	14/05/2019	18,676.20	22/06/2020	11,305.82	20/05/2021	19,233.29
20/05/2017	8,992.23	17/05/2018	12,500.10	15/05/2019	9,927.37	23/06/2020	8,877.86	21/05/2021	24,817.61
22/05/2017	8,113.57	18/05/2018	9,591.61	16/05/2019	17,716.84	24/06/2020	14,972.32	22/05/2021	16,347.58
23/05/2017	15,322.53	19/05/2018	5,457.33	17/05/2019	14,228.86	25/06/2020	8,506.85	23/05/2021	3,963.85
24/05/2017	12,275.04	20/05/2018	562.93	18/05/2019	5,508.96	26/06/2020	9,253.10	24/05/2021	17,831.22
25/05/2017	13,431.59	21/05/2018	19,230.50	19/05/2019	104.30	27/06/2020	6,458.87	25/05/2021	16,197.56
26/05/2017	16,121.65	22/05/2018	7,904.81	20/05/2019	12,005.05	30/06/2020	7,409.80	26/05/2021	17,887.69
27/05/2017	11,469.21	23/05/2018	17,938.62	21/05/2019	14,766.60	01/07/2020	6,143.46	27/05/2021	31,782.31
29/05/2017	18,453.30	24/05/2018	12,269.31	22/05/2019	13,598.26	02/07/2020	18,114.01	28/05/2021	25,885.42
30/05/2017	9,878.25	25/05/2018	22,382.35	23/05/2019	8,560.54	03/07/2020	9,025.73	29/05/2021	5,776.38

31/05/2017	12,198.09	26/05/2018	8,397.45	24/05/2019	20,273.35	04/07/2020	13,012.95	31/05/2021	11,626.63
01/06/2017	19,349.65	27/05/2018	117.55	25/05/2019	14,684.06	06/07/2020	9,130.67	01/06/2021	11,316.44
02/06/2017	7,516.13	28/05/2018	10,672.44	27/05/2019	19,891.06	07/07/2020	7,454.17	02/06/2021	12,591.71
03/06/2017	6,460.89	29/05/2018	28,295.19	28/05/2019	20,077.82	08/07/2020	9,774.34	03/06/2021	19,130.47
05/06/2017	13,747.48	30/05/2018	16,424.88	29/05/2019	16,580.96	09/07/2020	14,543.52	04/06/2021	10,073.09
06/06/2017	15,433.76	31/05/2018	24,684.66	30/05/2019	40,492.27	10/07/2020	10,901.05	05/06/2021	4,128.98
07/06/2017	22,734.12	01/06/2018	7,487.08	31/05/2019	10,488.76	11/07/2020	11,571.08	07/06/2021	7,960.21
08/06/2017	15,745.06	02/06/2018	3,371.20	01/06/2019	5,739.12	13/07/2020	7,057.36	08/06/2021	16,651.02
09/06/2017	13,167.97	03/06/2018	605.15	03/06/2019	15,893.26	14/07/2020	8,957.33	09/06/2021	16,883.58
10/06/2017	4,916.75	04/06/2018	10,207.18	04/06/2019	12,479.08	15/07/2020	8,706.09	10/06/2021	12,419.68
12/06/2017	15,815.11	05/06/2018	16,246.05	05/06/2019	8,823.06	16/07/2020	10,078.88	11/06/2021	12,822.35
13/06/2017	28,298.01	06/06/2018	12,663.99	06/06/2019	15,141.27	17/07/2020	17,899.20	12/06/2021	17,157.00
14/06/2017	17,982.86	07/06/2018	11,965.86	07/06/2019	17,203.57	18/07/2020	3,014.30	14/06/2021	12,567.04
15/06/2017	20,256.37	08/06/2018	14,536.36	08/06/2019	5,452.25	20/07/2020	6,306.31	15/06/2021	26,450.08
16/06/2017	11,679.08	09/06/2018	9,052.95	09/06/2019	597.22	21/07/2020	10,067.03	16/06/2021	14,403.32
17/06/2017	3,907.19	10/06/2018	143.48	10/06/2019	11,436.49	22/07/2020	7,145.56	17/06/2021	19,350.45
19/06/2017	8,502.32	11/06/2018	12,382.36	11/06/2019	11,714.17	23/07/2020	5,993.19	18/06/2021	14,951.33
20/06/2017	9,612.02	12/06/2018	13,545.23	12/06/2019	11,133.93	24/07/2020	18,103.93	19/06/2021	10,361.29
21/06/2017	8,720.10	13/06/2018	12,348.49	13/06/2019	16,250.58	25/07/2020	6,707.61	21/06/2021	14,154.42

22/06/2017	17,603.13	14/06/2018	24,181.56	14/06/2019	10,466.37	27/07/2020	5,873.11	22/06/2021	13,104.22
23/06/2017	10,750.25	15/06/2018	18,604.21	15/06/2019	2,342.97	28/07/2020	2.98	23/06/2021	17,874.17
24/06/2017	9,866.22	16/06/2018	2,541.31	16/06/2019	2,603.68	29/07/2020	8,606.10	24/06/2021	23,987.81
26/06/2017	17,010.56	18/06/2018	7,376.31	17/06/2019	12,557.12	30/07/2020	12,474.36	25/06/2021	18,035.16
27/06/2017	12,289.35	19/06/2018	7,474.42	18/06/2019	11,698.14	31/07/2020	11,426.17	26/06/2021	5,051.38
28/06/2017	16,764.87	20/06/2018	8,701.91	19/06/2019	28,486.51	01/08/2020	4,464.75	28/06/2021	15,775.03
29/06/2017	424.96	21/06/2018	9,959.63	20/06/2019	9,294.01	03/08/2020	4,393.62	29/06/2021	294.20
30/06/2017	35,036.00	22/06/2018	13,876.76	21/06/2019	17,093.17	04/08/2020	9,046.12	30/06/2021	42,416.20
01/07/2017	2,139.80	23/06/2018	9,213.60	22/06/2019	3,560.41	05/08/2020	16,432.58	01/07/2021	12,436.47
03/07/2017	8,608.71	25/06/2018	7,934.29	23/06/2019	0.00	06/08/2020	7,216.44	02/07/2021	12,945.35
04/07/2017	13,655.17	26/06/2018	11,105.25	24/06/2019	11,649.14	07/08/2020	10,490.87	03/07/2021	4,888.25
05/07/2017	16,948.25	27/06/2018	15,074.05	25/06/2019	22,045.52	08/08/2020	4,469.64	05/07/2021	11,482.53
06/07/2017	14,352.50	28/06/2018	29,810.78	26/06/2019	23,488.98	10/08/2020	7,404.35	06/07/2021	26,562.37
07/07/2017	19,403.86	29/06/2018	1,222.06	27/06/2019	15,439.16	11/08/2020	7,031.96	07/07/2021	17,715.69
08/07/2017	1,967.71	30/06/2018	10,675.68	28/06/2019	20,329.26	12/08/2020	9,192.90	08/07/2021	14,770.72
10/07/2017	8,782.71	02/07/2018	9,131.29	29/06/2019	4,317.85	13/08/2020	7,713.36	09/07/2021	24,798.54
11/07/2017	2,654.14	03/07/2018	7,140.19	30/06/2019	83.83	14/08/2020	10,835.07	10/07/2021	4,823.09
12/07/2017	915.05	04/07/2018	16,117.75	01/07/2019	1,790.69	15/08/2020	2,988.66	11/07/2021	468.12
13/07/2017	215.73	05/07/2018	7,459.54	02/07/2019	8,081.69	17/08/2020	7,480.69	12/07/2021	18,925.98

14/07/2017	10,150.44	06/07/2018	5,484.75	03/07/2019	7,421.48	18/08/2020	4,950.69	13/07/2021	12,980.78
15/07/2017	1,185.92	07/07/2018	5,857.79	04/07/2019	15,228.06	19/08/2020	7,859.33	14/07/2021	18,926.60
17/07/2017	41,205.63	09/07/2018	4,105.54	05/07/2019	24,060.88	20/08/2020	11,878.95	15/07/2021	24,547.67
18/07/2017	22,313.79	10/07/2018	15,287.11	06/07/2019	2,699.91	21/08/2020	10,344.89	16/07/2021	17,535.15
19/07/2017	24,332.80	11/07/2018	6,897.04	08/07/2019	10,024.50	22/08/2020	8,693.15	17/07/2021	15,702.74
20/07/2017	10,095.92	12/07/2018	13,578.09	09/07/2019	12,642.69	24/08/2020	5,011.70	19/07/2021	19,094.14
21/07/2017	17,918.85	13/07/2018	12,626.89	10/07/2019	13,861.73	25/08/2020	14,153.56	20/07/2021	28,386.76
22/07/2017	6,526.16	14/07/2018	1,721.05	11/07/2019	9,953.72	26/08/2020	9,320.16	21/07/2021	41,455.80
24/07/2017	19,889.46	16/07/2018	11,504.22	12/07/2019	15,379.58	27/08/2020	10,627.89	22/07/2021	17,384.25
25/07/2017	19,602.76	17/07/2018	12,112.82	13/07/2019	9,537.66	28/08/2020	8,869.82	23/07/2021	27,305.21
26/07/2017	5,724.04	18/07/2018	18,810.74	15/07/2019	15,456.90	29/08/2020	3,197.94	24/07/2021	11,937.52
27/07/2017	118.14	19/07/2018	13,562.78	16/07/2019	13,258.76	31/08/2020	17,952.86	26/07/2021	29,530.69
31/07/2017	7,206.64	20/07/2018	5,435.04	17/07/2019	13,914.94	01/09/2020	8,740.13	27/07/2021	37,213.32
01/08/2017	4,055.72	21/07/2018	5,669.83	18/07/2019	12,310.51	02/09/2020	10,551.72	30/07/2021	15,828.27
02/08/2017	5,578.43	23/07/2018	12,075.93	19/07/2019	10,311.81	03/09/2020	12,623.64	31/07/2021	17,768.50
03/08/2017	4,220.80	24/07/2018	11,686.76	20/07/2019	4,244.56	04/09/2020	7,774.00	02/08/2021	8,387.11
04/08/2017	4,738.91	25/07/2018	9,191.60	22/07/2019	12,230.23	05/09/2020	6,530.08	03/08/2021	20,454.61
05/08/2017	1,262.98	26/07/2018	16,712.87	23/07/2019	16,348.93	07/09/2020	6,705.96	04/08/2021	18,547.44
07/08/2017	7,699.12	27/07/2018	8,519.00	24/07/2019	7,891.80	08/09/2020	6,083.21	05/08/2021	18,661.71

08/08/2017	7,486.82	30/07/2018	6,168.07	25/07/2019	8,052.38	09/09/2020	9,044.38	06/08/2021	16,306.44
09/08/2017	7,222.40	31/07/2018	18,907.16	26/07/2019	9,349.46	10/09/2020	6,922.35	07/08/2021	13,546.55
10/08/2017	8,382.31	01/08/2018	12,732.87	27/07/2019	3,289.89	11/09/2020	10,146.17	09/08/2021	16,803.67
11/08/2017	7,187.94	02/08/2018	7,044.04	30/07/2019	13,923.21	12/09/2020	6,289.01	10/08/2021	20,166.42
12/08/2017	1,253.51	03/08/2018	12,635.99	31/07/2019	9,678.36	14/09/2020	8,398.96	11/08/2021	18,061.79
14/08/2017	5,201.68	04/08/2018	4,418.73	01/08/2019	15,070.03	15/09/2020	10,285.44	12/08/2021	27,447.49
15/08/2017	10,017.05	06/08/2018	8,573.03	02/08/2019	9,733.90	16/09/2020	15,533.37	13/08/2021	31,165.64
16/08/2017	7,051.22	07/08/2018	7,397.53	03/08/2019	3,907.48	17/09/2020	6,713.73	14/08/2021	8,567.58
17/08/2017	5,511.53	08/08/2018	11,535.84	04/08/2019	0.00	18/09/2020	7,647.25	16/08/2021	14,068.97
18/08/2017	9,470.54	09/08/2018	11,668.10	05/08/2019	11,758.71	19/09/2020	6,309.67	17/08/2021	18,039.65
19/08/2017	5,168.06	10/08/2018	13,996.00	06/08/2019	11,617.62	21/09/2020	10,417.27	18/08/2021	22,416.76
21/08/2017	12,763.53	11/08/2018	4,679.09	07/08/2019	20,993.68	22/09/2020	8,525.66	19/08/2021	16,427.30
22/08/2017	8,701.28	13/08/2018	10,875.84	08/08/2019	15,252.81	23/09/2020	14,542.54	20/08/2021	25,798.10
23/08/2017	7,037.96	14/08/2018	10,171.38	09/08/2019	8,944.64	24/09/2020	15,779.68	21/08/2021	25,790.71
24/08/2017	8,956.10	15/08/2018	14,643.23	10/08/2019	1,768.85	25/09/2020	16,222.69	23/08/2021	17,578.62
25/08/2017	16,216.24	16/08/2018	14,500.39	12/08/2019	5,878.36	26/09/2020	2,778.56	24/08/2021	19,869.87
26/08/2017	2,628.02	17/08/2018	10,353.06	13/08/2019	9,136.25	28/09/2020	11,223.31	25/08/2021	19,589.22
28/08/2017	8,611.00	18/08/2018	4,305.20	14/08/2019	7,217.92	29/09/2020	19,401.88	26/08/2021	25,111.79
29/08/2017	5,707.80	20/08/2018	7,732.57	15/08/2019	9,470.80	30/09/2020	15,352.25	27/08/2021	14,820.95

30/08/2017	701.30	21/08/2018	9,158.80	16/08/2019	10,305.29	01/10/2020	11,173.05	28/08/2021	16,353.60
31/08/2017	16,878.15	22/08/2018	8,702.08	17/08/2019	1,841.62	02/10/2020	9,845.28	31/08/2021	23,237.69
01/09/2017	6,325.94	23/08/2018	10,684.45	19/08/2019	5,656.11	03/10/2020	4,458.01	01/09/2021	18,780.56
02/09/2017	65.92	24/08/2018	31,614.44	20/08/2019	18,928.48	05/10/2020	6,753.72	02/09/2021	14,746.26
04/09/2017	6,900.64	25/08/2018	4,763.85	21/08/2019	5,418.15	06/10/2020	13,757.10	03/09/2021	21,489.40
05/09/2017	8,710.88	27/08/2018	6,978.04	22/08/2019	12,170.90	07/10/2020	16,455.41	04/09/2021	15,998.08
06/09/2017	6,013.21	28/08/2018	7,485.46	23/08/2019	14,631.29	08/10/2020	5,930.93	06/09/2021	16,475.58
07/09/2017	4,783.17	29/08/2018	20,045.18	24/08/2019	6,628.67	09/10/2020	9,469.37	07/09/2021	11,786.22
08/09/2017	10,828.51	30/08/2018	1,282.73	26/08/2019	14,294.66	10/10/2020	7,963.11	08/09/2021	23,459.49
09/09/2017	5,595.57	31/08/2018	10,247.50	27/08/2019	14,597.94	12/10/2020	5,923.72	09/09/2021	15,146.75
11/09/2017	10,828.37	01/09/2018	443.71	28/08/2019	20,208.00	13/10/2020	7,394.36	10/09/2021	22,831.53
12/09/2017	11,265.47	03/09/2018	12,434.74	29/08/2019	8,613.97	14/10/2020	15,799.98	11/09/2021	19,837.85
13/09/2017	17,547.91	04/09/2018	7,984.49	30/08/2019	3,143.94	15/10/2020	13,385.45	13/09/2021	21,794.41
14/09/2017	8,635.95	05/09/2018	4,627.06	31/08/2019	11,615.36	16/10/2020	9,002.31	14/09/2021	16,625.03
15/09/2017	10,962.86	06/09/2018	8,738.03	02/09/2019	8,417.30	17/10/2020	4,514.77	15/09/2021	19,828.11
16/09/2017	2,274.25	07/09/2018	10,875.65	03/09/2019	14,448.60	19/10/2020	16,908.92	16/09/2021	19,164.23
18/09/2017	12,315.14	08/09/2018	2,665.82	04/09/2019	12,587.30	20/10/2020	8,946.33	17/09/2021	30,442.47
19/09/2017	11,218.69	10/09/2018	8,345.72	05/09/2019	14,007.14	21/10/2020	11,894.42	18/09/2021	10,916.04
20/09/2017	8,333.06	11/09/2018	6,065.68	06/09/2019	6,628.80	22/10/2020	12,057.68	20/09/2021	9,026.93

21/09/2017	17,797.25	12/09/2018	18,118.58	07/09/2019	3,798.26	23/10/2020	7,133.13	21/09/2021	22,331.15
22/09/2017	10,617.94	13/09/2018	8,680.39	09/09/2019	8,647.34	24/10/2020	8,520.44	22/09/2021	26,389.45
23/09/2017	7,298.89	14/09/2018	4,883.22	10/09/2019	11,085.09	26/10/2020	9,161.24	23/09/2021	18,292.25
25/09/2017	8,317.43	15/09/2018	3,433.92	11/09/2019	15,593.86	27/10/2020	16,204.17		
26/09/2017	18,085.57	17/09/2018	23,590.54	12/09/2019	11,025.93	28/10/2020	15,469.17		
27/09/2017	17,546.80	18/09/2018	11,288.79	13/09/2019	16,508.52	29/10/2020	9,907.61		
28/09/2017	14,929.20	19/09/2018	8,214.85	14/09/2019	3,074.42	30/10/2020	25,116.23		
29/09/2017	10,962.39	20/09/2018	10,752.24	16/09/2019	10,821.96	31/10/2020	7,962.96		
30/09/2017	10,412.14	21/09/2018	19,773.94	17/09/2019	20,536.81	02/11/2020	6,416.35		
01/10/2017	129.12	22/09/2018	7,588.07	18/09/2019	19,863.22	03/11/2020	11,431.16		
02/10/2017	15,230.65	24/09/2018	8,956.09	19/09/2019	17,495.90	04/11/2020	10,279.56		
03/10/2017	5,867.49	25/09/2018	9,992.34	20/09/2019	5,477.65	05/11/2020	9,617.35		
04/10/2017	18,502.56	26/09/2018	12,794.30	21/09/2019	4,607.09	06/11/2020	10,062.24		
05/10/2017	6,937.96	27/09/2018	25,995.81	23/09/2019	16,264.64	07/11/2020	3,705.64		
06/10/2017	9,636.81	28/09/2018	20,275.56	24/09/2019	15,108.38	09/11/2020	5,889.84		
07/10/2017	5,341.97	29/09/2018	18,933.09	25/09/2019	20,650.63	10/11/2020	7,238.03		
08/10/2017	75.97	01/10/2018	6,279.05	26/09/2019	18,228.85	11/11/2020	9,998.07		
09/10/2017	7,161.57	02/10/2018	10,602.95	27/09/2019	19,949.23	12/11/2020	8,157.14		
10/10/2017	16,935.79	03/10/2018	12,958.97	28/09/2019	8,476.71	13/11/2020	12,234.44		

11/10/2017	4,242.88	04/10/2018	7,770.62	30/09/2019	21,619.61	14/11/2020	11,246.88
12/10/2017	11,821.94	05/10/2018	10,237.56	01/10/2019	11,050.37	16/11/2020	12,403.72
13/10/2017	9,097.80	06/10/2018	4,301.31	02/10/2019	16,253.05	17/11/2020	11,307.38
14/10/2017	4,214.48	08/10/2018	83.78	03/10/2019	13,065.31	18/11/2020	15,075.32
16/10/2017	5,874.75	09/10/2018	8,814.89	04/10/2019	14,408.64	19/11/2020	10,972.53
17/10/2017	6,373.40	10/10/2018	5,412.26	05/10/2019	6,745.79	20/11/2020	10,778.07
18/10/2017	5,393.26	11/10/2018	15,097.97	07/10/2019	19,557.64	21/11/2020	9,559.50
19/10/2017	12,063.59	12/10/2018	11,327.04	08/10/2019	2,820.45	23/11/2020	11,173.03
20/10/2017	10,830.15	13/10/2018	2,185.00	09/10/2019	21,414.78	24/11/2020	11,216.65
21/10/2017	4,948.92	15/10/2018	16,098.99	10/10/2019	20,174.66	25/11/2020	10,081.64
23/10/2017	14,094.65	16/10/2018	6,552.61	11/10/2019	23,574.56	26/11/2020	11,117.40
24/10/2017	11,853.46	17/10/2018	8,308.40	12/10/2019	5,290.05	27/11/2020	14,913.07
25/10/2017	18,115.59	18/10/2018	15,300.32	13/10/2019	123.23	28/11/2020	9,987.85
26/10/2017	15,375.64	19/10/2018	12,483.88	14/10/2019	5,721.13	30/11/2020	15,622.52
27/10/2017	16,093.25	20/10/2018	5,220.16	15/10/2019	13,224.55	01/12/2020	19,432.98
28/10/2017	9,561.19	22/10/2018	11,180.10	16/10/2019	12,415.26	02/12/2020	10,678.08
30/10/2017	11,668.03	23/10/2018	12,842.00	17/10/2019	9,276.19	03/12/2020	24,252.98
31/10/2017	22,906.46	24/10/2018	17,001.99	18/10/2019	16,653.08	04/12/2020	16,995.32
01/11/2017	841.79	25/10/2018	13,923.43	19/10/2019	3,325.59	05/12/2020	9,840.16

02/11/2017	18,476.53	26/10/2018	18,599.60	20/10/2019	0.00	07/12/2020	9,466.49
03/11/2017	7,032.36	27/10/2018	7,117.93	21/10/2019	17,426.75	08/12/2020	672.24
04/11/2017	5,185.96	29/10/2018	17,170.54	22/10/2019	14,261.78	09/12/2020	16,888.89
05/11/2017	120.13	30/10/2018	14,194.41	23/10/2019	21,395.53	10/12/2020	17,852.85
06/11/2017	15,053.23	31/10/2018	32,160.63	24/10/2019	21,307.61	11/12/2020	18,438.32
07/11/2017	6,543.62	01/11/2018	2,701.01	25/10/2019	16,692.85	12/12/2020	4,276.81
08/11/2017	23,503.01	02/11/2018	13,996.38	26/10/2019	4,678.41	14/12/2020	13,263.19
09/11/2017	9,884.19	03/11/2018	6,937.70	27/10/2019	0.00	15/12/2020	17,552.77
10/11/2017	13,669.94	05/11/2018	7,438.86	28/10/2019	16,433.52	16/12/2020	13,277.82
11/11/2017	3,269.45	06/11/2018	17,864.89	29/10/2019	10,292.63	17/12/2020	11,923.67
13/11/2017	11,993.56	07/11/2018	19,366.43	30/10/2019	45,836.05	18/12/2020	6,581.96
14/11/2017	13,852.13	08/11/2018	17,397.67	31/10/2019	29,208.05	19/12/2020	10,863.38
15/11/2017	22,047.94	09/11/2018	15,155.17	01/11/2019	2,200.02	21/12/2020	17,269.34
16/11/2017	9,028.86	10/11/2018	10,329.18	02/11/2019	4,446.57	22/12/2020	17,193.77
17/11/2017	16,738.85	12/11/2018	15,164.74	04/11/2019	14,736.44	23/12/2020	21,218.13
18/11/2017	4,993.55	13/11/2018	8,828.90	05/11/2019	7,918.10	24/12/2020	3,459.79
20/11/2017	10,559.66	14/11/2018	17,419.46	06/11/2019	11,638.22	28/12/2020	12,064.69
21/11/2017	12,235.92	15/11/2018	16,780.89	07/11/2019	19,458.26	29/12/2020	15,650.18
22/11/2017	25,814.18	16/11/2018	12,127.28	08/11/2019	15,783.49	30/12/2020	16,584.58

23/11/2017	9,315.81	17/11/2018	7,884.07	09/11/2019	1,859.65	31/12/2020	7,450.68
24/11/2017	27,527.16	19/11/2018	17,538.39	11/11/2019	5,745.82		
25/11/2017	2,193.17	20/11/2018	13,828.39	12/11/2019	8,839.57		
27/11/2017	19,679.25	21/11/2018	8,046.36	13/11/2019	22,330.72		
28/11/2017	17,155.13	22/11/2018	13,441.60	14/11/2019	12,488.33		
29/11/2017	16,886.08	23/11/2018	16,924.11	15/11/2019	17,548.54		
30/11/2017	14,342.87	24/11/2018	8,974.85	16/11/2019	4,960.00		
01/12/2017	7,890.84	26/11/2018	8,935.24	17/11/2019	0.00		
02/12/2017	2,304.71	27/11/2018	17,681.83	18/11/2019	18,690.83		
04/12/2017	7,137.70	28/11/2018	11,636.24	19/11/2019	18,918.27		
05/12/2017	9,645.98	29/11/2018	29,002.71	20/11/2019	34,633.48		
06/12/2017	25,529.70	30/11/2018	27,429.12	21/11/2019	25,385.97		
07/12/2017	24,227.67	01/12/2018	1,837.63	22/11/2019	24,070.05		
08/12/2017	1,870.95	03/12/2018	9,762.99	23/11/2019	986.19		
09/12/2017	6,126.03	04/12/2018	14,527.18	25/11/2019	14,395.66		
11/12/2017	10,638.76	05/12/2018	19,083.08	26/11/2019	8,412.23		
12/12/2017	12,681.89	06/12/2018	13,959.39	27/11/2019	21,750.64		
13/12/2017	14,441.99	07/12/2018	6,970.47	28/11/2019	8,001.84		
14/12/2017	13,809.00	08/12/2018	1,274.97	29/11/2019	13,300.90		

15/12/2017	12,707.22	10/12/2018	26,379.39	30/11/2019	3,687.60
16/12/2017	17,892.33	11/12/2018	7,997.18	02/12/2019	11,259.36
18/12/2017	8,827.85	12/12/2018	9,799.36	03/12/2019	10,652.09
19/12/2017	21,024.17	13/12/2018	6,433.51	04/12/2019	39,488.90
20/12/2017	11,486.39	14/12/2018	17,229.12	05/12/2019	15,116.48
21/12/2017	17,046.81	15/12/2018	8,488.06	06/12/2019	8,466.71
22/12/2017	31,938.40	17/12/2018	7,095.25	07/12/2019	4,455.34
23/12/2017	15,437.75	18/12/2018	20,035.88	09/12/2019	13,162.80
26/12/2017	6,105.65	19/12/2018	16,043.64	10/12/2019	13,203.56
27/12/2017	10,507.26	20/12/2018	13,145.07	11/12/2019	19,022.24
28/12/2017	11,116.72	21/12/2018	20,010.64	12/12/2019	24,112.01
29/12/2017	7,973.87	22/12/2018	12,043.95	13/12/2019	13,835.93
30/12/2017	1,840.13	26/12/2018	7,123.61	14/12/2019	7,367.39
		27/12/2018	16,457.50	16/12/2019	16,165.84
		28/12/2018	15,868.86	17/12/2019	11,435.63
		29/12/2018	6,704.96	18/12/2019	11,453.87
		31/12/2018	5,682.49	19/12/2019	19,997.85
				20/12/2019	14,127.35
				21/12/2019	10,860.57

22/12/2019	0.00
24/12/2019	6,369.53
26/12/2019	18,215.81
27/12/2019	23,455.65
28/12/2019	7,451.02
30/12/2019	23,509.65
31/12/2019	15,962.45

Nota. Fuente, elaboración propia.

3.3.3.7. Limpieza de datos

En este paso, para la limpieza de datos se seleccionaron siete (06) criterios de inclusión y siete (06) criterios de exclusión que permitan contar con los datos adecuados que posteriormente puedan ser probados y entrenados. Dichos criterios se conceptualizan a continuación:

Tabla 20

Criterios de inclusión y exclusión para limpieza de datos

Criterio	Descripción
Inclusión	El dato sí es necesario para ser probado y entrenado.
	El dato sí es relevante para su posterior transformación.
	El dato sí debe ser parte del grupo de datos suficientes para entrenar el modelo.
	El dato no es nulo o vacío.
	El dato sí es de tipo transaccional.
	El dato sí es correcto para el modelo propuesto según artículo científico.
Exclusión	El dato no es necesario para ser probado y entrenado.
	El dato no es relevante para su posterior transformación.
	El dato no debe ser parte del grupo de datos suficientes para entrenar el modelo.
	El dato sí es nulo o vacío.
	El dato no es de tipo transaccional.
	El dato no es correcto para el modelo propuesto según artículo científico.

Nota. Fuente, elaboración propia.

En la limpieza de datos, se definió que los datos nulos, erróneos y/o extraños se eliminarían.

Tabla 21

Datos de Valores nulos y extraños

FECHA	VALOR ACUMULADO POR DIA
02/01/2017	156.66
03/01/2017	4781.79
04/01/2017	7563.97
05/01/2017	11155.98
06/01/2017	8664.01
07/01/2017	12925.96
08/01/2017	NaN
09/01/2017	4337.7
10/01/2017	13804.76
11/01/2017	11185.96
12/01/2017	11195.48
13/01/2017	16397.61
14/01/2017	4246.08
15/01/2017	NaN
16/01/2017	8238.41
17/01/2017	10452.31
18/01/2017	11278.46
19/01/2017	6118.21

20/01/2017	30602.32
21/01/2017	6079.67
22/01/2017	NaN
23/01/2017	20326.68
24/01/2017	18359.18
25/01/2017	16406.53
26/01/2017	11323.71
27/01/2017	17864.56
28/01/2017	1414.71
29/01/2017	NaN
30/01/2017	11060.99
31/01/2017	21468.54

Fuente, Elaboración propia

3.3.3.8. Transformación de los datos

En esta actividad, los datos a emplearse que previamente fueron filtrados y son transformados realizando una descripción detallada de las columnas (variables) y filas seleccionadas, se realizaron los siguientes pasos:

1.- Se transformaron los valores de ventas que están expresados en soles a valores de (-1 a 1), Para lo cual se realizó la transformación de estos valores desde Python, convirtiéndolos en valores de peso equivalentes para el modelo.

Figura 42

Script conversión de pesos de ventas de los 7 días de predicción en Python

```
Crea el dataset con la estructura necesario para los input de nuestro modelo
de predicción. Y retorna el scaler usado en nuestros datos para posteriores
predicciones.
"""
PASOS = 7
scaler, data, scaled = build_feature(INTER_PATH + "inter_train.csv", PASOS, PROCC_FILE)#
convert series to supervised learning
def series_to_supervised(data, n_in=1, n_out=1, dropnan=True):
    n_vars = 1 if type(data) is list else data.shape[1]
    df = pd.DataFrame(data)
    cols, names = list(), list()
    # input sequence (t-n, ... t-1)
    for i in range(n_in, 0, -1):
        cols.append(df.shift(i))
        names += [('var%d(t-%d)' % (j+1, i)) for j in range(n_vars)]
    # forecast sequence (t, t+1, ... t+n)
    for i in range(0, n_out):
        cols.append(df.shift(-i))
        if i == 0:
            names += [('var%d(t)' % (j+1)) for j in range(n_vars)]
        else:
            names += [('var%d(t+%d)' % (j+1, i)) for j in range(n_vars)]
    # put it all together
    agg = pd.concat(cols, axis=1)
    agg.columns = names
    # drop rows with NaN values
    if dropnan:
        agg.dropna(inplace=True)
    return agg
def build_features_from(input, scaler, pasos, start_date, end_date):
    df = pd.read_csv(input).set_index("FECHA")
    df.index = pd.to_datetime(df.index)
    ultimosDias = df[start_date:end_date]
    values = ultimosDias.values.astype('float32')
    # normalize features
    values = values.reshape(-1, 1) # esto lo hacemos porque tenemos 1 sola dimension
    scaled = scaler.fit_transform(values)
    reframed = series_to_supervised(scaled, pasos, 1)
    reframed.drop(reframed.columns[[7]], axis=1, inplace=True)
    values = reframed.values
    x_test = values[6:, :]
    x_test = x_test.reshape((x_test.shape[0], 1, x_test.shape[1]))
    return x_test
```

Nota. Fuente, elaboración propia.

3.3.3.9. Desarrollo del Modelo

En esta etapa definiremos el modelo y los algoritmos a utilizar en la predicción de las ventas, como estrategia se tomó la serie temporal y se convirtió en un “problema de tipo supervisado”, para poder alimentar la red neuronal y entrenarla con backpropagation, teniendo en cuenta las entradas y las salidas para entrenar al modelo. En el modelo se tomó los 7 días previos para “obtener” el octavo día. Se definió:

- Entradas: “7 columnas” que representan las ventas en unidades de los 7 días anteriores.
- Salida: El valor del “8vo día”. Es decir, las ventas de ese día.
- Se usó como entradas las columnas encabezadas como $var1(t-7)$ a $(t-1)$ y como salida (lo que sería el valor “Y” de la función) será el $var1(t)$ la última columna.

Para poder entrenar y validar el modelo se establecieron los datos de Test y Validación, como estrategia se usaron el total de registro del dataset después de agruparlos por días, menos los últimos 29 días de ventas y para la validación se utilizó los registros de los últimos 29 días de ventas.

Tabla 22*Datos para el test de validación*

ITEM	FECHA	VALOR VENTA
1	20/08/2021	25798.10
2	21/08/2021	25790.71
3	23/08/2021	17578.62
4	24/08/2021	19869.87
5	25/08/2021	19589.22
6	26/08/2021	25111.79
7	27/08/2021	14820.95
8	28/08/2021	16353.60
9	31/08/2021	23237.69
10	1/09/2021	18780.56
11	2/09/2021	14746.26
12	3/09/2021	21489.40
13	4/09/2021	15998.08
14	6/09/2021	16475.58
15	7/09/2021	11786.22
16	8/09/2021	23459.49
17	9/09/2021	15146.75
18	10/09/2021	22831.53
19	11/09/2021	19837.85
20	13/09/2021	21794.41
21	14/09/2021	16625.03
22	15/09/2021	19828.11
23	16/09/2021	19164.23
24	17/09/2021	30442.47
25	18/09/2021	10916.04
26	20/09/2021	9026.93
27	21/09/2021	22331.15
28	22/09/2021	26389.45
29	23/09/2021	18292.25

Nota. Fuente, elaboración propia.

Se entrenó al Modelo de Red Neuronal normal FeedForward, definiendo la iteración a 70 épocas (iteraciones), utilizando el data set de test, también se realizó la validación con los datos de validación.

Finalmente se obtuvo los datos de predicción de la próxima semana del último día de venta de la data set total.

Se hicieron reuniones con la gerencia de ventas y se le explicó el modelo que se emplearía y nos dieron la aprobación para la implementación.

Se definió que al final del proceso de generación de modelos se dispondrá de tres tipos de información que pueden utilizarse en la toma de decisiones del análisis:

- Configuración de parámetros incluye las notas que ha tomado sobre los parámetros que producen los mejores resultados.
- Los modelos reales producidos.
- Descripciones de resultados de modelos, incluyendo problemas de datos y rendimiento que hayan ocurrido durante la ejecución del modelo y exploración de los resultados

Una vez definido el modelo y los algoritmos a utilizar en la predicción de las ventas, se procedió a la implantación de los mismos en el lenguaje de programación Python, utilizando herramientas de apoyo de ciencia de datos. A continuación, se ejemplifica el código de la aplicación:

Cargar los Valores del DataSet

- **Dividimos en set de Entrenamiento y Test Validación**

Por rangos de fecha de nuestro dataset

Figura 43

Ejecución de procedimiento de set entrenamiento y validación

```

train_range=["2017-01-02", "2020-12-31"],
    test_range=["2021-01-02", "2021-09-23"],
    output_path=INTER_PATH)
def xlsx_to_csv(filepath, sheet_name, cols, output):
    """Lee un fichero xlsx, procesa y almacena en csv"""
    df = pd.read_excel(filepath, sheet_name=sheet_name,
usecols=cols)
    # Calculamos el total Venta por dia (Acumulados)
    df = df.groupby('FECHA').sum()
    print("\nDataFrame INFO:\n")
    print(df.info())
    print("\n")
    df.to_csv(output)

def split_train_test(filepath, train_range, test_range, output_path):
    df = pd.read_csv(filepath).set_index("FECHA")
    df.index = pd.to_datetime(df.index)
    df.index.sort_values()

```

Nota. Fuente, elaboración propia.

Creamos el Modelo de Red Neuronal en este informe se utilizará la red FEEDFORWARD. Utilizando el optimizador Adam.

Figura 44

Ejecución del SCRIP de la red feedforward utilizando el optimizador ADAM

```

"""# Creamos el Modelo de Red Neuronal
## Utilizaremos una Red "normal" Feedforward y el optimizador Adam y
definimos los datos a obtener MSE,MAE,MAPE
"""
def crear_modeloFF(pasos):
    model = Sequential()
    model.add(Dense(pasos, input_shape=(1,pasos), activation='tanh'))
    model.add(Flatten())
    model.add(Dense(1, activation='tanh'))
    model.compile(loss='mean_absolute_error', optimizer='sgd',
metrics=["mse", "mae", "mape"])
    model.summary()
    return model

```

Nota. Fuente, elaboración propia.

- **Entrenamiento nuestra máquina**

Se conlleva a realizar el entrenamiento de nuestro modelo de la red neuronal, ejecutando las declaraciones de 70 épocas para el ratio de aprendizaje y se mostrara cómo se comporta la máquina, la cual se muestra a continuación

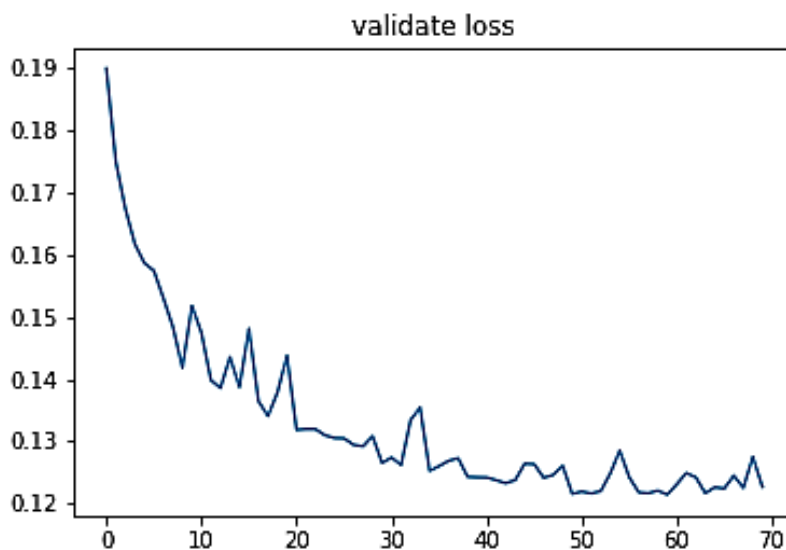
Figura 45

Ejecución del SCRIP entrenamiento de la red 70 épocas

```
""" Entrenamiento """
EPOCHS = 70
x_train, y_train, x_val, y_val, history, model = train(PROCC_FILE, PASOS,
EPOCHS, MODEL_FILE)
history.history['loss']
history.history['val_loss']
history.history['mse']
history.history['val_mse']
history.history['mae']
history.history['val_mae']
history.history['mape']
history.history['val_mape']
```

Figura 46

Ejecución del SCRIP para visualizar los resultados de entrenamiento de la red 70 épocas



Nota. Fuente, elaboración propia.

Tabla 23

Resultados de entrenamiento de las 70 épocas

Épocas	Tiempo	Error Absoluto Medio del test (MAE)	Error Cuadrático Medio del test (MSE)	Error Absoluto Relativo (MAPE)	Perdida de la validación (VAL_LOSS)	Error Cuadrático Medio de la Validación (VAL_MSE)	Error Absoluto Medio de la Validación del Test (VAL_MAE)	Error Absoluto Relativo de la Validación del Test (VAL_MAPE)
1	2s 5ms	0.2903	0.1681	85.159	0.2307	0.0726	0.2307	69.7516
2	0s 2ms	0.2187	0.0858	79.0905	0.2255	0.0699	0.2255	70.4048
3	0s 2ms	0.2053	0.0764	74.9087	0.2139	0.0628	0.2139	66.3607
4	0s 2ms	0.1957	0.0691	70.5049	0.2121	0.0622	0.2121	66.5999
5	0s 2ms	0.1899	0.0651	69.4091	0.2058	0.0592	0.2058	64.4616
6	0s 2ms	0.1859	0.0621	67.8523	0.1992	0.0557	0.1992	61.5888
7	0s 2ms	0.1826	0.0597	66.7913	0.2005	0.0557	0.2005	61.9727
8	0s 2ms	0.1802	0.0578	65.2557	0.1992	0.0548	0.1992	61.3232
9	0s 2ms	0.1784	0.0567	64.9922	0.2019	0.0555	0.2019	62.3416
10	0s 2ms	0.1768	0.056	64.1997	0.2009	0.0551	0.2009	61.8732
11	0s 2ms	0.1753	0.0552	64.2282	0.1884	0.0509	0.1884	56.173
12	0s 2ms	0.1746	0.055	63.8602	0.2006	0.0552	0.2006	61.8666
13	0s 2ms	0.1742	0.0546	63.8587	0.1911	0.0519	0.1911	57.6791
14	0s 2ms	0.1731	0.0542	62.8747	0.1971	0.0542	0.1971	60.6666
15	0s 2ms	0.1733	0.0543	64.0741	0.1802	0.0496	0.1802	51.9842
16	0s 2ms	0.1727	0.0538	62.1363	0.19	0.0519	0.19	57.5382

17	0s 2ms	0.173	0.0539	63.2869	0.1979	0.055	0.1979	61.4557
18	0s 2ms	0.173	0.0538	63.3771	0.1866	0.051	0.1866	56.3213
19	0s 2ms	0.1726	0.0538	63.2812	0.1834	0.0503	0.1834	54.7796
20	0s 2ms	0.173	0.0538	62.7747	0.1917	0.053	0.1917	59.0928
21	0s 2ms	0.1722	0.0539	63.3346	0.1864	0.0512	0.1864	56.6544
22	0s 2ms	0.1718	0.0533	63.0068	0.1904	0.0526	0.1904	58.9652
23	0s 2ms	0.1727	0.0541	63.404	0.1858	0.0512	0.1858	56.562
24	0s 2ms	0.1722	0.0536	63.5922	0.1894	0.0524	0.1894	58.3457
25	0s 2ms	0.172	0.0539	63.5941	0.1765	0.0491	0.1765	51.1621
26	0s 2ms	0.1723	0.0538	62.9149	0.1812	0.05	0.1812	54.4541
27	0s 2ms	0.1719	0.0536	63.1929	0.189	0.0523	0.189	58.4721
28	0s 2ms	0.1722	0.0537	63.0546	0.1949	0.0546	0.1949	61.1338
29	0s 2ms	0.172	0.0535	63.5446	0.1942	0.0544	0.1942	61.0391
30	0s 2ms	0.1719	0.0534	63.2293	0.1883	0.0521	0.1883	58.5169
31	0s 2ms	0.1718	0.0534	63.1531	0.1834	0.0504	0.1834	56.2364
32	0s 2ms	0.1717	0.0535	63.302	0.1931	0.0541	0.1931	60.7778
33	0s 2ms	0.1715	0.0533	63.109	0.1936	0.0543	0.1936	61.0502
34	0s 2ms	0.1717	0.0533	63.0888	0.1832	0.0504	0.1832	56.2513
35	0s 2ms	0.1711	0.0533	62.9122	0.1921	0.0538	0.1921	60.5733
36	0s 2ms	0.1717	0.0532	63.752	0.1865	0.0516	0.1865	57.9726
37	0s 2ms	0.1716	0.0532	63.2311	0.187	0.0519	0.187	58.2291
38	0s 2ms	0.1717	0.0534	63.4472	0.182	0.0501	0.182	55.5324
39	0s 2ms	0.1716	0.0535	63.0805	0.1813	0.0499	0.1813	54.9943
40	0s 2ms	0.1718	0.0532	62.8138	0.188	0.0523	0.188	58.4535

41	0s 2ms	0.1712	0.0533	63.2732	0.191	0.0535	0.191	59.9404
42	0s 2ms	0.1715	0.0532	62.7626	0.1828	0.0505	0.1828	55.8169
43	0s 2ms	0.1713	0.0531	62.4676	0.1949	0.055	0.1949	61.5355
44	0s 2ms	0.1713	0.0532	62.7673	0.1882	0.0526	0.1882	58.5062
45	0s 2ms	0.1715	0.0532	63.0288	0.1798	0.0496	0.1798	53.8299
46	0s 2ms	0.1711	0.0533	62.7886	0.1808	0.0498	0.1808	54.5655
47	0s 2ms	0.1713	0.0534	63.5073	0.1831	0.0506	0.1831	56.0008
48	0s 2ms	0.1709	0.0528	62.6549	0.1804	0.0497	0.1804	54.1653
49	0s 2ms	0.1708	0.0533	62.302	0.1859	0.0517	0.1859	57.3452
50	0s 2ms	0.171	0.0533	62.8234	0.1835	0.0508	0.1835	55.8467
51	0s 2ms	0.1716	0.0532	62.5141	0.1844	0.0511	0.1844	56.5075
52	0s 2ms	0.17	0.0528	62.3703	0.2025	0.0582	0.2025	64.5962
53	0s 2ms	0.1719	0.0534	63.5177	0.1849	0.0512	0.1849	56.8889
54	0s 2ms	0.1709	0.0531	63.0756	0.192	0.0539	0.192	60.1772
55	0s 2ms	0.1713	0.0532	62.711	0.1863	0.0519	0.1863	57.5851
56	0s 2ms	0.1706	0.0531	63.1811	0.1763	0.0492	0.1763	51.1549
57	0s 2ms	0.1708	0.053	62.4534	0.1824	0.0506	0.1824	55.3737
58	0s 2ms	0.1706	0.0531	62.722	0.1853	0.0516	0.1853	57.0722
59	0s 2ms	0.1711	0.0534	63.2658	0.188	0.0526	0.188	58.2723
60	0s 2ms	0.1704	0.0526	62.7584	0.1855	0.0517	0.1855	57.0926
61	0s 2ms	0.1704	0.0528	62.7781	0.1828	0.0506	0.1828	55.4139
62	0s 2ms	0.1701	0.053	63.1224	0.1789	0.0495	0.1789	53.1721
63	0s 2ms	0.1704	0.0529	62.7876	0.196	0.0554	0.196	61.9539
64	0s 2ms	0.1707	0.0531	63.2085	0.1859	0.0517	0.1859	57.4077

65	0s 2ms	0.1704	0.0529	62.7125	0.2006	0.0573	0.2006	63.6599
66	0s 2ms	0.1708	0.053	63.0047	0.1924	0.0541	0.1924	60.3294
67	0s 3ms	0.1698	0.0528	63.0386	0.1862	0.0517	0.1862	57.5083
68	0s 3ms	0.1702	0.0528	62.697	0.1886	0.0528	0.1886	58.4309
69	0s 2ms	0.1701	0.0529	62.7065	0.1871	0.0522	0.1871	57.5548
70	0s 2ms	0.1698	0.0527	62.6204	0.1868	0.0521	0.1868	57.2382

Nota. Fuente, elaboración propia.

- **Resultados del test de Validación de los últimos 29 días**

Se procedió a realizar el entrenamiento de nuestro modelo, ejecutando los 29 últimos días del dataset desde el día 20/08/2021 al 23/09/2021 obteniendo resultados reales y pronosticados, la cual se muestra en la tabla 24.

Figura 47

Ejecución del SCRIP para visualizar los resultados de pronóstico

```

"""
Predicción y gráficos del conjuntos de datos TESTING
"""
start = "2021-08-20"
end = "2021-09-23"
print(f"\nTesting de {start} a {end}")
values = df[start:end].values.astype('float32')
y_true, y_pred = predict_test(model, scaler, values, PASOS)

```

Nota. Fuente, elaboración propia.

Tabla 24

Resultados de pronósticos del test de validación del dataset

DIAS	FECHA	Datos
		Reales Ventas
1	20/08/2021	25798.10
2	21/08/2021	25790.71
3	23/08/2021	17578.62
4	24/08/2021	19869.87
5	25/08/2021	19589.22
6	26/08/2021	25111.79
7	27/08/2021	14820.95
8	28/08/2021	16353.60
9	31/08/2021	23237.69
10	1/09/2021	18780.56
11	2/09/2021	14746.26

12	3/09/2021	21489.40
13	4/09/2021	15998.08
14	6/09/2021	16475.58
15	7/09/2021	11786.22
16	8/09/2021	23459.49
17	9/09/2021	15146.75
18	10/09/2021	22831.53
19	11/09/2021	19837.85
20	13/09/2021	21794.41
21	14/09/2021	16625.03
22	15/09/2021	19828.11
23	16/09/2021	19164.23
24	17/09/2021	30442.47
25	18/09/2021	10916.04
26	20/09/2021	9026.93
27	21/09/2021	22331.15
28	22/09/2021	26389.45
29	23/09/2021	18292.25

Nota. Fuente, elaboración propia.

Tabla 25*Valores del test de validación de los pesos asignados en los valores reales y resultados predictivos*

DIAS	real	pred	mean_ absolute_error	mse	mae	mape
1	-0.42406392	-0.48075941	0.05669549	0.00321438	0.05669549	0.13369563
2	-0.51942843	-0.50349236	-0.01593607	0.00025396	0.01593607	0.03068002
3	-0.5507864	-0.54907107	-0.00171533	2.94E-06	0.00171533	0.00311432
4	-0.505246	-0.5496586	0.0444126	0.00197248	0.0444126	0.08790292
5	-0.52946615	-0.50327277	-0.02619338	0.00068609	0.02619338	0.0494713
6	-0.53166866	-0.53527135	0.00360269	1.30E-05	0.00360269	0.00677619
7	-0.42444962	-0.54475296	0.12030334	0.01447289	0.12030334	0.28343372
8	-0.4495166	-0.47235426	0.02283766	0.00052156	0.02283766	0.05080494
9	-0.22416633	-0.51407623	0.2899099	0.08404775	0.2899099	1.29328032
10	-0.33389324	-0.28756082	-0.04633242	0.00214669	0.04633242	0.13876417
11	-0.22692794	-0.49449176	0.26756382	0.0715904	0.26756382	1.17906951
12	-0.08673579	-0.28676325	0.20002747	0.04001099	0.20002747	2.30617001
13	-0.40658486	-0.24970792	-0.15687694	0.02461037	0.15687694	0.38584058
14	-0.77711886	-0.49230072	-0.28481814	0.08112137	0.28481814	0.36650525
15	-0.61240107	-0.64112365	0.02872258	0.00082499	0.02872258	0.04690159

16	-0.23112929	-0.47793022	0.24680093	0.0609107	0.24680093	1.06780464
17	-0.19674432	-0.15346704	-0.04327728	0.00187292	0.04327728	0.21996709
18	-0.49802357	-0.31280464	-0.18521893	0.03430605	0.18521893	0.37190796
19	-0.31825012	-0.61239523	0.29414511	0.08652134	0.29414511	0.92425765
20	-0.31792545	-0.39861256	0.08068711	0.00651041	0.08068711	0.25379254
21	-0.44284636	-0.3946127	-0.04823366	0.00232649	0.04823366	0.10891737
22	-0.59144074	-0.43527073	-0.15617001	0.02438907	0.15617001	0.26405015
23	-0.6149707	-0.6068064	-0.0081643	6.67E-05	0.0081643	0.01327592
24	-0.23731261	-0.54609793	0.30878532	0.09534838	0.30878532	1.30117537
25	-0.14968115	-0.20023005	0.0505489	0.00255519	0.0505489	0.33771051
26	-0.4544654	-0.30110955	-0.15335585	0.02351802	0.15335585	0.3374423
27	-0.38706946	-0.56870753	0.18163807	0.03299239	0.18163807	0.46926478
28	-0.28676414	-0.46344557	0.17668143	0.03121633	0.17668143	0.61612109
29	-0.4687273	-0.3462978	-0.1224295	0.01498898	0.1224295	0.26119558

Nota. Fuente, elaboración previa.

- **Preparación de datos para la Predicción**

Se realizará la preparación de datos del mes de setiembre 2021, prediciendo las ventas de la próxima semana.

Figura 48

Ejecución del SCRIP para visualizar los datos reales de los días 28/08/2021 al 23/09/2021

```

*****
df = pd.read_csv(INTER_FILE).set_index("FECHA")
df.index = pd.to_datetime(df.index)

ultimosDias = df['2021-09-01':'2021-09-23']
ultimosDias

```

Tabla 26

Valores reales de los últimos días del mes de setiembre del año 2021

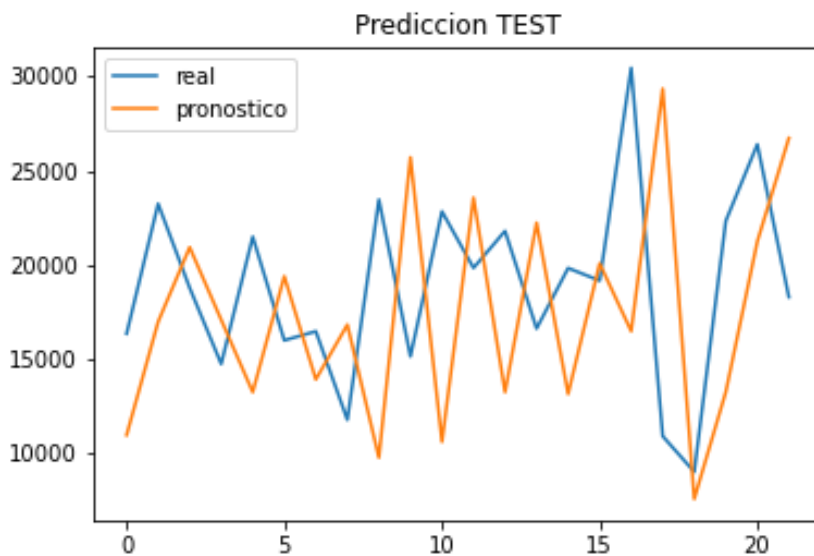
DIAS	FECHA	VALOR VENTA	
		REAL	PRONOSTICO
1	28/08/2021	16,353.60	10,975.89
2	31/08/2021	23,237.69	16,981.82
3	1/09/2021	18,780.56	20,931.30
4	2/09/2021	14,746.26	17,084.71
5	3/09/2021	21,489.40	13,254.38
6	4/09/2021	15,998.08	19,392.39
7	6/09/2021	16,475.58	13,928.44
8	7/09/2021	11,786.22	16,826.86
9	8/09/2021	23,459.49	9,780.91
10	9/09/2021	15,146.75	25,700.20
11	10/09/2021	22,831.53	10,619.50
12	11/09/2021	19,837.85	23,560.29
13	13/09/2021	21,794.41	13,253.38
14	14/09/2021	16,625.03	22,230.74

15	15/09/2021	19,828.11	13,155.61
16	16/09/2021	19,164.23	20,072.39
17	17/09/2021	30,442.47	16,484.05
18	18/09/2021	10,916.04	29,338.69
19	20/09/2021	9,026.93	7,589.45
20	21/09/2021	22,331.15	13,260.86
21	22/09/2021	26,389.45	21,241.62
22	23/09/2021	18,292.25	26,722.47

Nota. Fuente, elaboración propia.

Figura 49

Valores reales de los últimos días del mes de setiembre del año 2021



Nota. Fuente, elaboración propia.

- **Preparamos los datos para realizar el proceso de obtener la predicción de los 7 días**

Se realiza el proceso de transformación correspondiente de los días 16/09/2021 al 30/09/2021. Incluyendo los 7 días de predicción desde día 24/09/2021 al

30/09/2021 tomando como día inicial 23/09/2021. Para ello se ejecuta el siguiente script:

Figura 50

Ejecución del SCRIP para visualizar los pesos correspondientes de los días 16/09/2021 al 30/09/2021

```
Predicción acumulativa de 7 días
"""
print("*****")
print("Predicción 7 Dias")
print("*****")
N = 7
e_dt = np.datetime64("2021-09-23")
s_dt = e_dt - PASOS
start = np.datetime_as_string(s_dt)
end = np.datetime_as_string(e_dt)
values = df[s_dt:e_dt].values.astype('float32')
n_pred, acc = predict_n_days(model, scaler, values, PASOS, N)
plot_accumulative_prediction(acc, N)
print(n_pred)
```

Nota. Fuente, elaboración propia.

Tabla 27

Valores de la predicción de 7 días

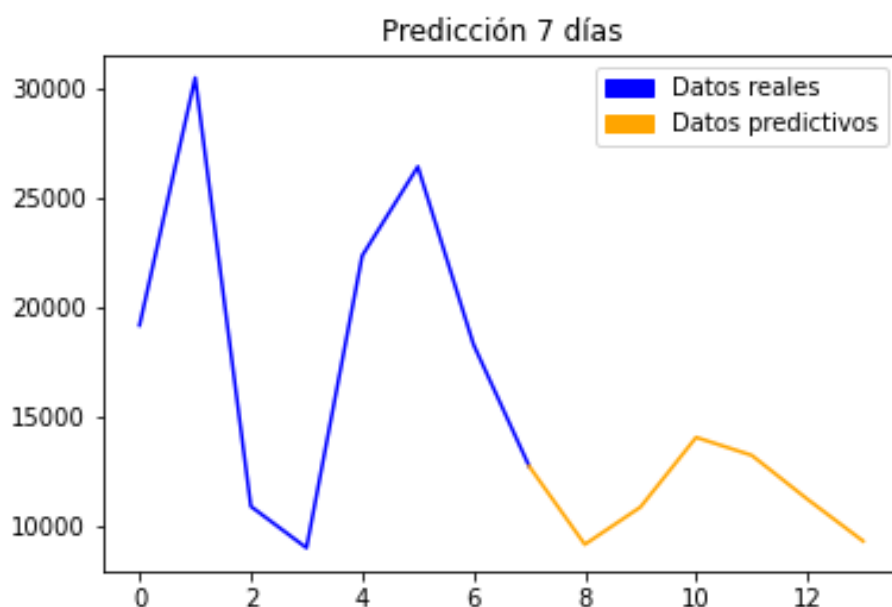
FECHA	VALOR VENTA REAL	PREDICCION
16/09/2021	19,164.23	19,164.23
17/09/2021	30,442.47	30,442.47
18/09/2021	10,916.04	10,916.04
20/09/2021	9,026.93	9,026.93
21/09/2021	22,331.15	22,331.15
22/09/2021	26,389.45	26,389.45
23/09/2021	18,292.25	18,292.25
24/09/2021	-	14,739.48
25/09/2021	-	11,279.25

26/09/2021	-	12,456.31
27/09/2021	-	15,296.98
28/09/2021	-	14,878.19
29/09/2021	-	13,088.73
30/09/2021	-	11,221.94

Nota. Fuente, elaboración propia.

Figura 51

Valores de la predicción de 7 días



Nota. Fuente, elaboración propia.

3.3.4. Resultados obtenidos en cuanto a la predicción del modelo implementado previamente

3.3.4.1. Evaluación entrenamiento de la red

Para nuestro entrenamiento de la red neuronal y evaluar la efectividad de nuestro modelo se realizó la declaración de 70 épocas para el ratio del aprendizaje. En nuestro caso aplicamos la función de pérdida o, Loos function,

que es aquella función que valúa la desviación entre los valores reales de las observaciones empleadas durante el aprendizaje y las predicciones desplegadas por la red neuronal. Al realizar el proceso nos dio como resultados como muestra la tabla siguiente:

Tabla 28

Valores de la Pérdida del test entrenamiento

Épocas	Pérdida del test LOSS
1	0.2307
2	0.2255
3	0.2139
4	0.2121
5	0.2058
6	0.1992
7	0.2005
8	0.1992
9	0.2019
10	0.2009
11	0.1884
12	0.2006
13	0.1911
14	0.1971
15	0.1802
16	0.19
17	0.1979
18	0.1866
19	0.1834
20	0.1917
21	0.1864
22	0.1904
23	0.1858
24	0.1894

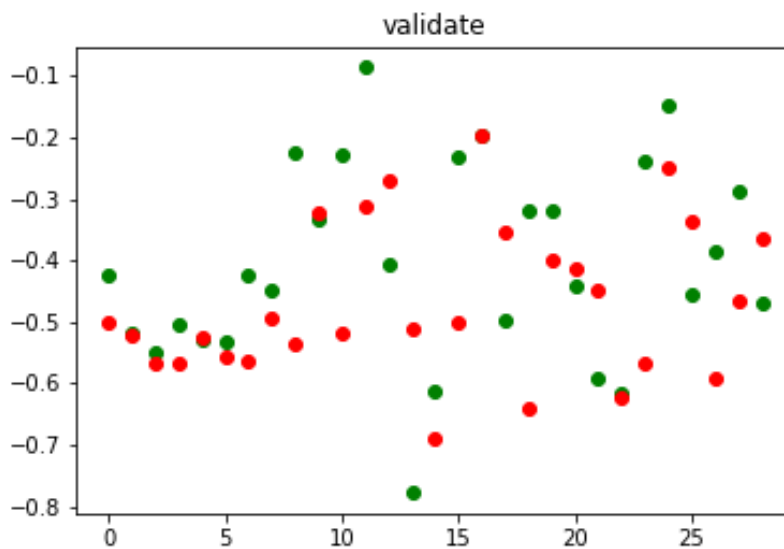
25	0.1765
26	0.1812
27	0.189
28	0.1949
29	0.1942
30	0.1883
31	0.1834
32	0.1931
33	0.1936
34	0.1832
35	0.1921
36	0.1865
37	0.187
38	0.182
39	0.1813
40	0.188
41	0.191
42	0.1828
43	0.1949
44	0.1882
45	0.1798
46	0.1808
47	0.1831
48	0.1804
49	0.1859
50	0.1835
51	0.1844
52	0.2025
53	0.1849
54	0.192
55	0.1863
56	0.1763

57	0.1824
58	0.1853
59	0.188
60	0.1855
61	0.1828
62	0.1789
63	0.196
64	0.1859
65	0.2006
66	0.1924
67	0.1862
68	0.1886
69	0.1871
70	0.1868

Nota. Fuente, elaboración propia.

Figura 52

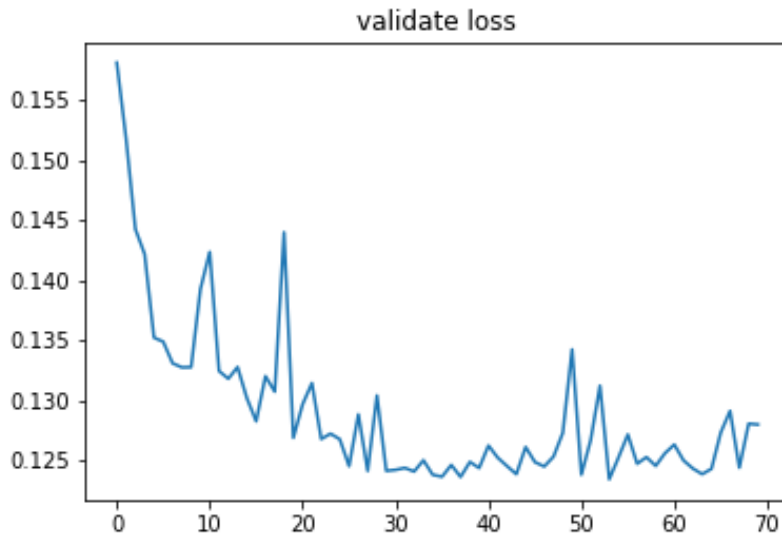
Diagrama de dispersión de 70 épocas



Nota. Fuente, elaboración propia.

Figura 53

Diagrama de dispersión de validate loss



Nota. Fuente, elaboración propia.

Como evidenciamos el modelo está aprendiendo como le vamos enseñando época por época en tal sentido se demuestra que, en la época 1 y 2 tenemos un rango de pérdida una pérdida de 0.2307 a 0.2255 también observamos que en la época 20 hasta la época 30 observamos que las pérdidas disminuyen de 0.1917 hasta 0.1834 en la época 60 evidenciamos que el modelo sigue mejorando en 0.1855 hasta la época 70 con un valor de 0.1868. Siendo los valores efectivos con una tendencia a mejorar.

3.3.4.2. Evaluación del test de entrenamiento

En el estudio realizado para lograr el modelo predictivo se realizó el análisis de nuestro data set (Datos totales=1451). Los cuales eran los datos disponibles y consistentes para la implementación del modelo. Para realizar nuestro entrenamiento se realizó lo siguiente:

- Consideración de Datos óptimos por año según muestra la tabla 3
- Del total de datos óptimos (1451) se optó por tomar 22 datos para realizar el entrenamiento del modelo.

Tabla 29

Valores de los datos óptimos por año del data set

AÑO	CANT. DATOS ÓPTIMOS
2017	312
2018	316
2019	318
2020	279
2021	226
TOTAL	1451
DATOS DE ENTRENAMIENTO	22

Nota. Fuente, elaboración previa.

Al realizar nuestro test de validación se trabajó con los últimos 22 días del año 2021 desde el día 28/08/2021 al 23/09/2021 obteniendo la predicción de estos días y comparándolos con los datos reales obteniendo el error del pronóstico como se conceptualiza a continuación:

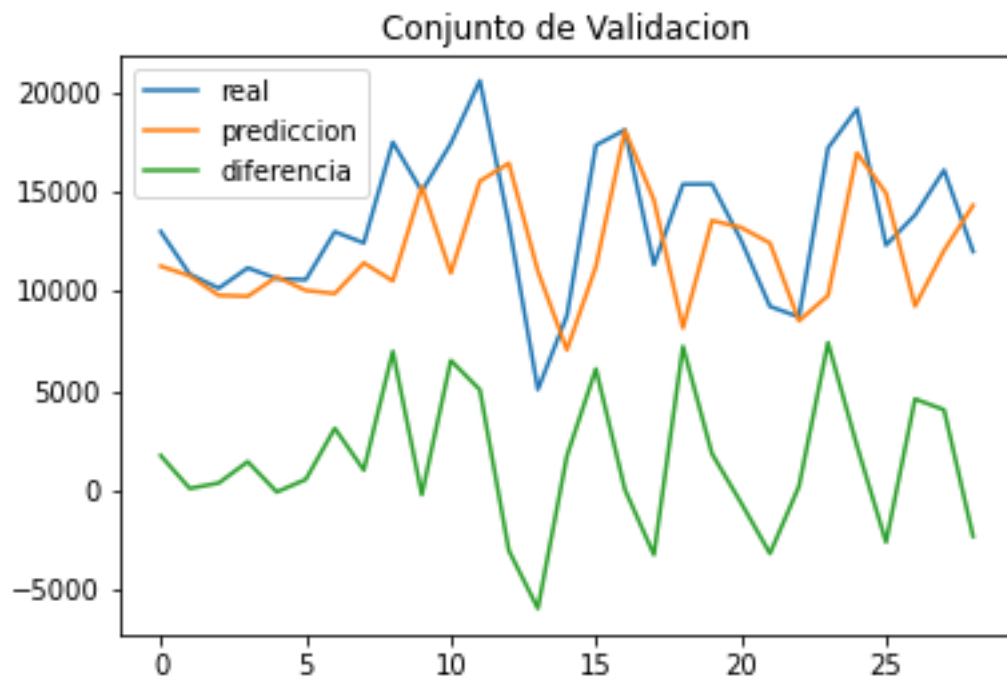
Tabla 30*Valores de los datos predictivos de los últimos 22 días del test de validación*

DIAS	FECHA	VALOR		
		VENTA REAL	PRONOSTICO	DIFERENCIA
1	28/08/2021	16,353.60	10,975.89	5,377.71
2	31/08/2021	23,237.69	16,981.82	6,255.87
3	1/09/2021	18,780.56	20,931.30	-2,150.74
4	2/09/2021	14,746.26	17,084.71	-2,338.45
5	3/09/2021	21,489.40	13,254.38	8,235.03
6	4/09/2021	15,998.08	19,392.39	-3,394.31
7	6/09/2021	16,475.58	13,928.44	2,547.15
8	7/09/2021	11,786.22	16,826.86	-5,040.64
9	8/09/2021	23,459.49	9,780.91	13,678.58
10	9/09/2021	15,146.75	25,700.20	-10,553.45
11	10/09/2021	22,831.53	10,619.50	12,212.03
12	11/09/2021	19,837.85	23,560.29	-3,722.44
13	13/09/2021	21,794.41	13,253.38	8,541.03
14	14/09/2021	16,625.03	22,230.74	-5,605.71
15	15/09/2021	19,828.11	13,155.61	6,672.50
16	16/09/2021	19,164.23	20,072.39	-908.15
17	17/09/2021	30,442.47	16,484.05	13,958.43
18	18/09/2021	10,916.04	29,338.69	-18,422.65
19	20/09/2021	9,026.93	7,589.45	1,437.48
20	21/09/2021	22,331.15	13,260.86	9,070.29
21	22/09/2021	26,389.45	21,241.62	5,147.83
22	23/09/2021	18,292.25	26,722.47	-8,430.22

Nota. Fuente, elaboración propia.

Figura 54

Cuadro estadístico del pronóstico de ventas de los últimos 22 días para el test de entrenamiento del modelo



Nota. Fuente, elaboración propia.

Al realizar la evaluación de los resultados del test de entrenamiento de los últimos 22 días del año 2021 se puede concluir que los valores predictivos obtenidos por el modelo son efectivos por lo que existe una tendencia de forma horizontal según la línea de tendencia del valor predictivo como se muestra la Figura 54.

3.3.4.3. Evaluación del pronóstico de 7 días

Para poder obtener la predicción de los 7 días a partir del día 24/09/2021 al 30/09/2021 se analizó los datos reales de los días 16/09/2021 al 23/09/2021, para ver la tendencia de los valores de nuestro modelo al brindar los pronósticos de próximos 7 días como se conceptualiza en la Figura 51.

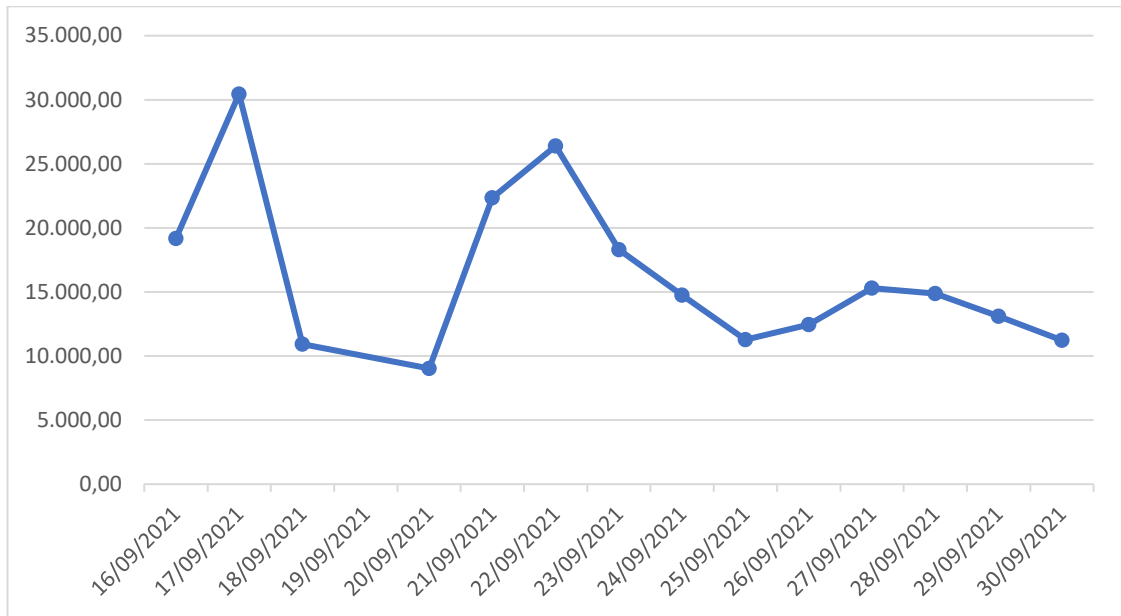
Tabla 31*Valores de los datos de los 7 días pronosticados*

FECHA	Pronostico Valor Venta
16/09/2021	19,164.23
17/09/2021	30,442.47
18/09/2021	10,916.04
20/09/2021	9,026.93
21/09/2021	22,331.15
22/09/2021	26,389.45
23/09/2021	18,292.25
24/09/2021	14,739.48
25/09/2021	11,279.25
26/09/2021	12,456.31
27/09/2021	15,296.98
28/09/2021	14,878.19
29/09/2021	13,088.73
30/09/2021	11,221.94

Nota. Fuente, elaboración propia.

Figura 55

Cuadro estadístico del pronóstico de ventas de los 7 días



Nota. Fuente, elaboración propia.

Al obtener los resultados como los valores predictivos de los días 24 al 30 de setiembre se verifico que existe una tendencia optima sin mostrar decadencia en los valores de venta en los próximos días como se muestra la **Figura 55**.

IV. CONCLUSIONES Y RECOMENDACIONES

4.1. Conclusiones

Se seleccionó al modelo de aprendizaje automático MLP-Backpropagation como el más empleado para la predicción de ventas de entre los modelos encontrados en la revisión sistemática de la literatura tales como Regresión Logística, Support Vector Machine y Naive Bayes.

Se diseñó el modelo predictivo para el proceso de ventas utilizando inteligencia de negocios y data Analytics, el cual quedó constituido por cuatro (04) procesos principales con el que se predijeron las ventas de la empresa textil caso de estudio de modo que, se lograron identificar las variables de las ventas, por tanto gracias al aporte de ingeniería se comprobó que, con datos pocos consistentes, las redes neuronales sí son de gran aporte para predecir las ventas a pesar de contar con data no válida.

Se implementó el modelo predictivo del proceso de ventas en la empresa caso de estudio, teniendo en cuenta los datos de las ventas diarias (del último día), para pronosticar las ventas de cada día tomando en consideración los datos históricos de los 7 días anteriores, realizando un análisis de regresión de series de tiempo haciendo el uso de información de desplazamiento histórica combinada con una red neuronal de siete capas de entrada siete capas ocultas y una capa de salida.

Se evaluaron los resultados de la predicción de ventas utilizando el modelo predictivo propuesto determinándose que, para poder obtener la predicción de los 7 días a partir del día 24/09/2021 al 30/09/2021 debían de analizarse los datos reales de los días 16/09/2021 al 23/09/2021, para ver la tendencia de los valores del modelo al brindar los pronósticos de próximos 7 días verificándose que, existe una tendencia optima sin mostrar decadencia en los valores de venta en los próximos días.

4.2. Recomendaciones

Para lograr determinar una predicción de ventas de forma óptima y rápida se sugiere utilizar el modelo propuesto en las pequeñas y medianas empresas entrenando y probando datos específicos del área de estudio.

Se exhorta a la gerencia general que aplique el uso de esta herramienta tecnológica y a la vez apoye con los recursos necesarios como de emplear de forma permanente con un científico de datos permitiéndole el análisis de las informaciones en tiempo real y estar a la vanguardia empresarial del mercado textil teniendo como premisas el crecimiento, la rentabilidad y su sostenibilidad.

Al obtener buenos resultados a partir del modelo propuesto de aprendizaje automático con las variables propuestas se propone que se incluya otras variables adicionales como clientes, productos para poder tener mayor proyección de las ventas por segmentación del mercado.

Se sugiere a las empresas peruanas de software que en su desarrollo de los aplicativos integren el modelo aplicado de redes neuronales en su desarrollo de sistemas y estos se relacionen con mayores alcances en cuanto análisis de información que conlleven a tener informes de proyecciones tales como los pronósticos así ser soporte a la toma de decisiones de las empresas.

REFERENCIAS

- Abdellatif, M., Shaaban, E., & Abu-Raya, K. (2019). Egyptian Case Study-Sales forecasting model for automotive section. *2019 International Conference on Smart Applications, Communications and Networking (SmartNets)* (págs. 1-6). Sharm el-Sheij: IEEE. Obtenido de <https://ieeexplore.ieee.org/document/9069751>
- Alnoukari, M., & Hanano, A. (2017). Integration of business intelligence with corporate strategic management. *Journal of Intelligence Studies in Business*, 7(2), 5-16. Obtenido de <https://journal.hh.se/index.php/JISIB/article/view/235>
- Alpana, B. (2002). *Reservoir Properties from Well Logs Using Neural Networks*. Obtenido de http://www.ipt.ntnu.no/~bjornu/Alpana_Bhatt.pdf .
- Álvarez, J. (2017). *La gestión empresarial de las Mypes y su relación con la competitividad de ventas en el emporio de gamarra, La Victoria, 2017*. Lima: Universidad César Vallejo. Obtenido de <https://repositorio.ucv.edu.pe/handle/20.500.12692/3055>
- Amable, M., Millones, R., & Checa, R. (2015). Calidad de software en las mypes productoras de software en Lima. *Memoria COMTEL 2015 - VII Congreso Internacional de Computación y Telecomunicaciones* (págs. 133-139). Lima, Perú: Fondo Editorial - Universidad Inca Garcilaso de la Vega. Obtenido de <http://www.comtel.pe/memoriacomtel/COMTEL2015.pdf>
- Angelo, J. (2006). *Business Intelligence - A new technology can analyze data at amazing speed. So why is higher ed slow to adopt?* Obtenido de University Business: www.universitybusiness.com
- ATRIA Innovation. (22 de Octubre de 2019). *Qué son las redes neuronales y sus funciones*. Obtenido de ATRIA: <https://www.atriainnovation.com/que-son-las-redes-neuronales-y-sus-funciones/>
- Baum, D. (2006). *The face of Intelligence*. Obtenido de Oracle Magazine: <https://www.oracle.com/technetwork/oramag/magazine/home/index.html>
- Big Data Marketer. (1 de Octubre de 2018). *¿Qué es el Análisis Predictivo?* Obtenido de Big Data Social: <http://www.bigdata-social.com/que-es-el->

- CRISP-DM consortium. Obtenido de <https://www.semanticscholar.org/paper/54bad20bbc7938991bf34f86dde0babfbd2d5a72>
- Cheng, X., Yang, B., & Lin, Z. (2018). A random forest learning assisted “divide and conquer” approach for peptide conformation search. *Scientific Reports*, 8(1), 87-96. Obtenido de <https://www.nature.com/articles/s41598-018-27167-w>
- Cruz, M., & Rúa, J. (2018). Surgimiento y desarrollo del método Delphi: una perspectiva cuantitativa. *Biblios*, 1(71), 90-107. Obtenido de http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1562-47302018000200007
- Dairu, X., & Shilong, Z. (2021). Machine Learning Model for Sales Forecasting by Using XGBoost. *2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE)* (págs. 480-483). Guangzhou: IEEE. Obtenido de <https://ieeexplore.ieee.org/document/9342304>
- Dairu, X., & Shilong, Z. (2021). Machine Learning Model for Sales Forecasting by Using XGBoost. *2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE)* (págs. 480-483). Guangzhou: IEEE. Obtenido de <https://ieeexplore.ieee.org/document/9342304>
- Davenport, T., & Harris, J. (2007). The Dark Side of Customer Analytics. *Harvard business review*, 85(5), 37-48. Obtenido de <https://www.pomsmeetings.org/ConfPapers/052/052-0002.pdf>
- Decide. (2021). *Analítica Avanzada*. Obtenido de Decide Soluciones: <https://decidesoluciones.es/analitica-avanzada/>
- Diario Gestión. (6 de Febrero de 2020). *Negocios 4.0: ¿cómo será el profesional que necesitarán las empresas en el futuro?* Obtenido de Redacción Gestión: <https://gestion.pe/economia/negocios-40-como-sera-el-profesional-que-necesitaran-las-empresas-en-el-futuro-fotos-nndc-noticia/>

- Díaz, J., Navares, R., Linares, C., & Aznarte, J. (2018). Comparing ARIMA and computational intelligence methods to forecast daily hospital admissions due to circulatory and respiratory causes in Madrid. *Stochastic Environmental Research and Risk Assessment*, 32(1), 2849-2859. Obtenido de <https://link.springer.com/article/10.1007/s00477-018-1519-z>
- Digital House. (15 de Septiembre de 2021). *Data Analytics: qué es y para qué se utiliza*. Obtenido de Digital House: <https://www.digitalhouse.com/ar/blog/data-analytics-se-utiliza-data-analytics>
- Dong, Y., Xie, K., Bohan, Z., & Lin, L. (2021). A Machine Learning Model for Product Fraud Detection Based On SVM. *2021 2nd International Conference on Education, Knowledge and Information Management (ICEKIM)* (págs. 385-388). Xiamen: IEEE. Obtenido de <https://ieeexplore.ieee.org/abstract/document/9479632>
- Dong, Y., Xie, K., Bohan, Z., & Lin, L. (2021). A Machine Learning Model for Product Fraud Detection Based On SVM. *2021 2nd International Conference on Education, Knowledge and Information Management (ICEKIM)* (págs. 385-388). Xiamen: IEEE. Obtenido de <https://ieeexplore.ieee.org/document/9479632>
- Espinosa, F. (2020). *Sistemas de información para la gestión de la empresa*. Talca: Universidad de Talca. Obtenido de [docplayer.es: https://docplayer.es/1624245-Sistemas-de-informacion-para-la-gestion-de-la-empresa-fernando-espinosa-fuentes.html](https://docplayer.es/1624245-Sistemas-de-informacion-para-la-gestion-de-la-empresa-fernando-espinosa-fuentes.html)
- Fan, Y. (2021). Prediction of A-share trend based on machine learning technologies. *2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA)* (págs. 282-287). Shenyang: IEEE. Obtenido de <https://ieeexplore.ieee.org/document/9362611>
- Flórez, R. &. (2008). Las Redes Neuronales Artificiales. En R. &. Flórez, Flórez, R., & Fernández, J *Fundamentos teóricos y aplicaciones prácticas*. España: Netbiblo, S. L.

- Gallego-Nicasio, J., Rodríguez, A., Mínguez, J., & Jiménez, F. (2018). Modelos ARIMA para la predicción del gasto conjunto de oxígeno de vuelo y otros gases en el Ejército del Aire. *Sanidad Militar*, 74(4), 223-229. Obtenido de https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1887-85712018000400223&lng=pt&nrm=iso
- Gestal, M. (2014). *Introducción a las redes de neuronas artificiales*.
- Giri, C., Thomassey, S., Balkow, J., & Zeng, X. (2019). Forecasting New Apparel Sales Using Deep Learning and Nonlinear Neural Network Regression. *2019 International Conference on Engineering, Science, and Industrial Applications (ICESI)* (págs. 1-6). Tokio: IEEE. Obtenido de <https://ieeexplore.ieee.org/document/8863024>
- González, H., Calleja, A., Gómez, G., Rodríguez, A., & López de Lacalle, L. (6 de Abril de 2021). *Los conceptos de Machine Learning y Deep Learning en la industria*. Obtenido de Interempresas: <https://www.interempresas.net/MetalMecanica/Articulos/347471-Los-conceptos-de-Machine-Learning-y-Deep-Learning-en-la-industria.html>
- González, J. (2015). *¿Cómo entender el Teorema de Bayes en forma simple?* Obtenido de docirs.cl: https://www.docirs.cl/entender_teorema_de_bayes_simple.asp
- Gonzalez, J. (Diciembre de 2015). *¿Cómo entender el Teorema de Bayes en forma simple?* Obtenido de DocIRS: https://www.docirs.cl/entender_teorema_de_bayes_simple.asp
- Gorry, G., & Morton, M. (1971). *A framework for management information systems* (Primera ed.). Massachusetts: Massachusetts Institute of Technology.
- Gruman, G. (2 de Abril de 2007). *Rethinking business intelligence*. Obtenido de IDG Communications Inc: <https://www.infoworld.com/article/2661363/rethinking-business-intelligence.html>
- Grupo Salesforce. (2021). *¿Qué es la inteligencia de negocios? Guía sobre la inteligencia de negocios y por qué es importante*. Obtenido de Tableau: <https://www.tableau.com/es-es/learn/articles/business-intelligence>

- Hernández, R., Fernández, C., & Baptista, M. (2014). *Metodología de la investigación científica*. México: McGRAW-HILL / INTERAMERICANA EDITORES,.
- Hernández-Orallo, J. (2021). CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories. *IEEE Transactions on Knowledge and Data Engineering*, 33(8), 3048-3061. Obtenido de <https://ieeexplore.ieee.org/abstract/document/8943998>
- Higuita, D. (20 de Agosto de 2016). *Analítica y productividad aplicada al ruteo de vehículos*. Obtenido de [grupobancolombia.com: https://www.grupobancolombia.com/wps/portal/empresas/capital-inteligente/tendencias/innovacion/analitica-y-productividad](https://www.grupobancolombia.com/wps/portal/empresas/capital-inteligente/tendencias/innovacion/analitica-y-productividad)
- Hofseth, L. (2018). Getting rigorous with scientific rigor. *Carcinogenesis*, 39(1), 21-25. Obtenido de <https://academic.oup.com/carcin/article/39/1/21/4100559>
- Imhoff, C., & Pettit, R. (2004). The Critical Shift to Flexible Business Intelligence: What Every Marketer Wants—And Needs—From Technology. *An intelligent solutions white paper*, 1(1), 1-19. Obtenido de http://download.101com.com/pub/tdwi/files/Critical_Shift_to_Flex_BI_Imhoff.pdf
- Jourdan, Z., Rainer, R., & Marshall, T. (2008). Business Intelligence: An Analysis of the Literature. *Information systems management*, 25(2), 121-131. Obtenido de <https://www.tandfonline.com/doi/abs/10.1080/10580530801941512>
- Keras. (2021). *Optimizers*. Obtenido de Comunidad Keras: <https://keras.io/api/optimizers/>
- Kingma, D., & Ba, J. (2017). Adam: a method for stochastic Optimization. *3rd International Conference on Learning Representations, ICLR 2015* (págs. 1-15). California: DBLP. Obtenido de <https://arxiv.org/abs/1412.6980>
- Kurnia, P. (2018). Business intelligence model to analyze social media information. *Procedia Computer Science*, 135(1), 5-14. Obtenido de <https://www.sciencedirect.com/science/article/pii/S1877050918314303>
- Lal, V. (2005). The Future of Business Intelligence from Hyperion. *Hyperion*, 1-12.

- Muñoz, G. (8 de Octubre de 2019). *La inspiración desbloquea el futuro*. Obtenido de ¿Dónde está Avinash cuando se le necesita?: <http://dóndeestaavinashcuandoseleneesita.com/analisis-predictivo/>
- Nunnari, G., & Nunnari, V. (2017). Forecasting Monthly Sales Retail Time Series: A Case Study. *2017 IEEE 19th Conference on Business Informatics (CBI)* (págs. 1-6). Salónica: IEEE. Obtenido de <https://ieeexplore.ieee.org/document/8010700>
- Olavsrud, T. (29 de Mayo de 2020). *Decision support systems: Sifting data for better business decisions*. Obtenido de CIO: <https://www.cio.com/article/3545813/decision-support-systems-sifting-data-for-better-business-decisions.html>
- Pavlyshenko, B. (2019). Machine-learning models for sales time series forecasting. *Data*, 4(1), 1-11. Obtenido de <https://www.mdpi.com/2306-5729/4/1/15>
- Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2008). Systematic Mapping Studies in Software Engineering. *12th International Conference on Evaluation and Assessment in Software Engineering (EASE)* (págs. 1-10). Swindon: Electronic Workshops in Computing (eWiC). Obtenido de <https://www.scienceopen.com/hosted-document?doi=10.14236/ewic/EASE2008.8>
- Pici, M. (2 de Julio de 2021). *8 Proven Sales Forecasting Methods for Greater Accuracy*. Obtenido de Sales Hacker: <https://www.saleshacker.com/sales-forecasting-methods/>
- Porter, M. (1980). Industry structure and competitive strategy: Keys to profitability. *Financial analysts journal*, 36(4), 30-41. Obtenido de <https://www.tandfonline.com/doi/abs/10.2469/faj.v36.n4.30?journalCode=ufaj20>
- Punam, K., Pamula, R., & Jain, P. (2018). A Two-Level Statistical Model for Big Mart Sales Prediction. *2018 International Conference on Computing, Power and Communication Technologies (GUCON)* (págs. 617-620). Greater Noida: IEEE. Obtenido de <https://ieeexplore.ieee.org/document/8675060>

- ReasonWhy. (31 de Enero de 2020). *Estas son las 10 tendencias que marcarán al consumidor en 2020*. Obtenido de ReasonWhy.: <https://www.reasonwhy.es/actualidad/tendencias-consumo-2020-consumidor-llyc>
- Rikhardsson, P., & Yigitbasioglu, O. (2018). Business intelligence & analytics in management accounting research: Status and future focus. *International Journal of Accounting Information Systems*, 29(1), 37-58. Obtenido de <https://www.sciencedirect.com/science/article/abs/pii/S1467089516300616>
- Salas, R. (2016). *Redes Neuronales Artificiales*. Valparaíso: Universidad de Valparaíso. Obtenido de <https://www.researchgate.net/publication/266882116>
- Sales Forecast of Chinese New Energy Vehicles Based on Wavelet and BP Neural Network. (2019). *2019 18th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES)* (págs. 141-144). Wuhan: IEEE. Obtenido de <https://ieeexplore.ieee.org/document/8920997>
- Tang, Z. (2014). Un algoritmo de aprendizaje supervisado por múltiples capas de redes neuronales. En Z. Tang, *Un algoritmo de aprendizaje supervisado por múltiples capas de redes neuronales* (págs. 1125 – 1142).
- Turban, E., Sharda, R., Aronsson, J., & King, D. (2008). *Business intelligence: A managerial approach* (Primera ed.). Indiana: Pearson Prentice Hall.
- Wang, C., Huang, R., Shen, J., & Liu, J. (2021). A Novel Lightweight Authentication Protocol for Emergency Vehicle Avoidance in VANETs. *IEEE Internet of Things Journal*, 1(1), 99-101. Obtenido de <https://www.researchgate.net/publication/350334232>
- Wong, T., & Yeh, S. (2019). Weighted Random Forests for Evaluating Financial Credit Risk. *Proceedings of Engineering and Technology Innovation* (págs. 1-9). Taipéi: Taiwan Association of Engineering and Technology Innovation: E-Journals. Obtenido de <https://core.ac.uk/download/pdf/228834303.pdf>

Zheng, Y., & Li, Y. (2018). Unmanned Retail's Distribution Strategy Based on Sales Forecasting. *2018 8th International Conference on Logistics, Informatics and Service Sciences (LISS)* (págs. 1-5). Toronto: IEEE. Obtenido de <https://ieeexplore.ieee.org/document/8593273>

ANEXOS

Anexo 1. Resolución de aprobación del proyecto de investigación.



FACULTAD DE INGENIERÍA, ARQUITECTURA Y URBANISMO
RESOLUCIÓN N° 0700-2021/FIAU-USS
Pimentel, 02 de Agosto de 2021

VISTO:

Las Actas de reunión N° 2306 - 2021, N° 1307 - 2021 y la N° 1507 - 2021 del Comité de investigación de la Escuela profesional de INGENIERÍA DE SISTEMAS remitida el 15 de Julio de 2021 mediante oficio N°0250-2021/FIAU-IS-USS de la Dirección de Escuela de INGENIERÍA DE SISTEMAS, y;

CONSIDERANDO:

Que, de conformidad con la Ley Universitaria N° 30220 en su artículo 48° que a letra dice: "La investigación constituye una función esencial y obligatoria de la universidad, que la fomenta y realiza, respondiendo a través de la producción de conocimiento y desarrollo de tecnologías a las necesidades de la sociedad, con especial énfasis en la realidad nacional. Los docentes, estudiantes y graduados participan en la actividad investigadora en su propia institución o en redes de investigación nacional o internacional, creadas por las instituciones universitarias públicas o privadas."

Que, de conformidad con el Reglamento de grados y títulos en su artículo 21° señala: "Los temas de trabajo de investigación, trabajo académico y tesis son aprobados por el Comité de Investigación y derivados a la Facultad o Escuela de Posgrado, según corresponda, para la emisión de la resolución respectiva. El periodo de vigencia de los mismos será de dos años, a partir de su aprobación. En caso un tema perdiera vigencia, el Comité de Investigación evaluará la ampliación de la misma.

Que, de conformidad con el Reglamento de grados y títulos en su artículo 24° señala: La tesis es un estudio que debe denotar rigurosidad metodológica, originalidad, relevancia social, utilidad teórica y/o práctica en el ámbito de la escuela profesional. Para el grado de doctor se requiere una tesis de máxima rigurosidad académica y de carácter original. Es individual para la obtención de un grado; es individual o en pares para obtener un título profesional. Asimismo, en su artículo 25° señala: "El tema debe responder a alguna de las líneas de investigación institucionales de la USS S.A.C."

Que, según documentos de Vistos el Comité de investigación de la Escuela profesional de INGENIERÍA DE SISTEMAS acuerda aprobar los temas de las Tesis a cargo de los egresados que se detallan en el anexo de la presente Resolución.

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

SE RESUELVE:

ARTÍCULO 1°: APROBAR, los **Temas de Tesis** perteneciente a la línea de investigación de **INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE**, a cargo de los **alumnos y egresados** del Programa de estudios de **INGENIERÍA DE SISTEMAS**, según se detalla en el anexo de la presente Resolución.

ARTÍCULO 2°: ESTABLECER, que la inscripción de los Temas de Tesis se realice a partir de emitida la presente resolución y tendrá una vigencia de dos (02) años.

ARTÍCULO 3°: DEJAR SIN EFECTO, toda Resolución emitida por la Facultad que se oponga a la presente Resolución.

REGÍSTRESE, COMUNÍQUESE Y ARCHÍVESE



Dr. Mario Fernando Ramos Moscol
Decano - Facultad de Ingeniería,
Arquitectura y Urbanismo
UNIVERSIDAD SEÑOR DE SIPÁN S.A.C.



Mag. Luis Roberto Larrea Colchaco
Secretario Académico (a)
Facultad de Ingeniería, Arquitectura y Urbanismo
UNIVERSIDAD SEÑOR DE SIPÁN S.A.C.

Cc: Interesado, Archivo

FACULTAD DE INGENIERÍA, ARQUITECTURA Y URBANISMO
RESOLUCIÓN N° 0700-2021/FIAU-USS
 Pimentel, 02 de Agosto de 2021

ANEXO

ESCUELA PROFESIONAL	APELLIDOS Y NOMBRES	TEMA DE TESIS	LÍNEA DE INVESTIGACIÓN
INGENIERÍA DE SISTEMAS	BECERRA SANCHEZ MIGUEL ANGEL	EVALUACION DE RENDIMIENTO DE PROTOCOLOS DE COMUNICACIÓN IOT EN EL MONITOREO DE LA CALIDAD DEL AIRE	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE
INGENIERÍA DE SISTEMAS	ÑAÑEZ PEREDO ALEXIS JOEL VELASCO CHERO JUNIOR	ADAPTACIÓN DE UNA METODOLOGÍA DE DESARROLLO ÁGIL DE SOFTWARE PARA MYPES PERUANAS DE DESARROLLO DE SOFTWARE	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE
INGENIERÍA DE SISTEMAS	VASQUEZ MUÑOZ DEXTRE ALEXANDER	DESARROLLO DE UN MÉTODO DE CONTEO AUTOMÁTICO DE CAPSIUM ANNUUM GROUP BASADO EN APRENDIZAJE PROFUNDO PARA LA ESTIMACIÓN DE SU PRODUCCIÓN	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE
INGENIERÍA DE SISTEMAS	CELIS SANCHEZ SUJEILY PEREZ ROJAS FRANKLIN ALEXI	MODELO DE EVALUACION DE PROCESOS TI CON BASE EN EL MARCO DE REFERENCIA COBIT 5 PAM. CASO DE ESTUDIO SIEMPRESOFT E.I.R.L. CHICLAYO	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE
INGENIERÍA DE SISTEMAS	PEREZ SILVA EDWIN WILDOR	RECONOCIMIENTO DE PLACAS VEHICULARES MEDIANTE VISION COMPUTACIONAL PARA MEJORAR EL ACCESO A UN PARQUEADERO	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE
INGENIERÍA DE SISTEMAS	LUQUE CONDORI BASILIO	EVALUACIÓN DE LA CALIDAD DE APLICACIONES GENERADAS POR BPM PARA LA GESTIÓN DE TRÁMITES EN UNA GERENCIA REGIONAL DE TRANSPORTE PERUANO	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE
INGENIERÍA DE SISTEMAS	BRENIS LLAGUENTO JULIO ANTONIO	COMPARACIÓN DE PROTOCOLOS DE AUTENTICACIÓN EN CONEXIONES DE REDES PRIVADAS VIRTUALES PARA USO EN TRABAJO REMOTO. CASO DEESTUDIO: MUNICIPALIDAD PROVINCIAL DE FERREÑAFE	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE
INGENIERÍA DE SISTEMAS	AGUILERA ALVARADO YEISSER FROILAN	EVALUACIÓN DEL PERFORMANCE EN LA COMUNICACIÓN DE APLICACIONES CON ARQUITECTURA DE SOFTWARE QUE UTILIZA SERVICIOS WEB EN EMPRESAS CON PRESENCIA NACIONAL	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE
INGENIERÍA DE SISTEMAS	JIMENEZ ESPINOZA LUIS ALBERTO	ANÁLISIS DE MÉTODOS DE RECONOCIMIENTO DE EXPRESIONES FACIALES	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE
INGENIERÍA DE SISTEMAS	CARREÑO GUERRERO SANTIAGO ANIBAL	MODELO PREDICTIVO DEL PROCESO DE VENTAS UTILIZANDO INTELIGENCIA DE NEGOCIOS Y DATA ANALYTICS EN LA EMPRESA CENTRO TEXTIL DE LA MATTA S.A.C	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE

FACULTAD DE INGENIERÍA, ARQUITECTURA Y URBANISMO
RESOLUCIÓN N° 0700-2021/FIAU-USS
Pimentel, 02 de Agosto de 2021

INGENIERÍA DE SISTEMAS	VILCHEZ SILVA OMAR JHONATHAN	ANÁLISIS COMPARATIVO DE ALGORITMO DEVITERBI Y BAM-WELCH DE RECONOCIMIENTO DE VOZ PARA LA IDENTIFICACIÓN DE PERSONAS	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE
INGENIERÍA DE SISTEMAS	FERNANDEZ SALDAÑA CARLOS ALDAIR IDROGO CORNEJO LEONCIO	DESARROLLO DE UN MÉTODO DE CLASIFICACIÓN AUTOMÁTICA DE CITRUS AURANTIFOLIA USANDO PROCESAMIENTO DIGITAL DE IMÁGENES	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE
INGENIERÍA DE SISTEMAS	SUCLUPE CHAPOÑAN MANUELA DEL ROSARIO	EVALUACIÓN DE LA CALIDAD DE UN SISTEMA DE INFORMACIÓN LOGÍSTICA DE CONTROL DEL PROCESO DE ALMACENAMIENTO DE UNA CONSTRUCTORA DE LA SELVA PERUANA	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE
INGENIERÍA DE SISTEMAS	PISFIL JAUREGUI JOSE LAURENT	IMPLEMENTACIÓN DE MICROSERVICIOS PARA MEJORAR LA INTEROPERABILIDAD DEL PROCESO DE CONSULTAS DE ARBITRIOS TRIBUTARIOS EN MUNICIPALIDADES	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE
INGENIERÍA DE SISTEMAS	CHIMPEN SERQUEN MERLINA JESSICA	DISEÑO DE UNA MESA DE AYUDA BASADO EN ITIL V3 Y BPM PARA EL AREA DE TI DE UNA MUNICIPALIDAD PROVINCIAL DE LAMBAYEQUE	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE
INGENIERÍA DE SISTEMAS	CIEZA CELIS JESUS ABELARDO OJEDA ROMERO ANTHONNY JHONATAN	EVALUACIÓN DEL DESEMPEÑO DE PROTOCOLOS DE SEGURIDAD DE REDES PARA COMBATIR VULNERABILIDADES EN REDES INALÁMBRICAS WI-FI	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE
INGENIERÍA DE SISTEMAS	FERNANDEZ RIOJA JUAN NICANOR RIOJA MESIA CHARLES SEGUNDO	IMPLEMENTACIÓN DE UN MODELO DE GESTIÓN BASADO EN ITIL PARA MEJORAR LAS ATENCIONES DE INCIDENCIAS DE TI EN UNA MUNICIPALIDAD DISTRITAL DE LA REGIÓN LAMBAYEQUE	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE
INGENIERÍA DE SISTEMAS	TEMOCHE GÓMEZ LENNIN BILLEY	MÉTODO PARA DETECTAR CON EFICIENCIA LOS ATAQUES CROSS-SITE SCRIPTING UTILIZANDO TÉCNICAS DE APRENDIZAJE AUTOMÁTICO	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE
INGENIERÍA DE SISTEMAS	BOBADILLA CAMPOS ROLANDO MARTÍN GUEVARA CHAMBERGO JHON DENNIS	IMPLEMENTACIÓN DE UNA METODOLOGÍA DE GESTIÓN DE RIESGOS AD HOC BASADA EN ESTÁNDARES INTERNACIONALES Y BUENAS PRÁCTICAS PARA UNA EMPRESA MANUFACTURERA PERUANA	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE
INGENIERÍA DE SISTEMAS	CALDERON YNOÑAN PAMELA DEL CARMEN PRIETO NEIRA FRANCK ALBERSON	DESARROLLO DE UN MÉTODO BAJO EL ENFOQUE ÁGIL Y DISEÑO CENTRADO EN USUARIO (DCU) PARA LA MEJORA DE EXPERIENCIA DE USUARIO (UX)	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE
INGENIERÍA DE SISTEMAS	CASTRO MEDINA MIGUEL ANGEL	IMPLEMENTACIÓN DE UN MODELO AD HOC DE GESTIÓN DE LA SEGURIDAD DE LA INFORMACIÓN PARA UNA EMPRESA EDITORA DE DIARIO REGIONAL PERUANO	INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE

Anexo 2. Carta de aceptación de la institución para la recolección de datos




La victoria 01 de Junio del 2021

Señor: Mg Ing, Victor Tuesta Monteza
Director de la escuela de ingeniería de sistemas
Universidad Señor de Sipan
Presente.

Asunto: Autorización al estudiante de ingeniería de sistemas para que realice trabajo de investigación de tesis profesional.

Estimados Señores: Les comunico que el estudiante del 10mo Ciclo de la escuela académica Profesional de ingeniería de sistemas, CARREÑO GUERRERO SANTIAGO ANIBAL, con código universitario 2151818827, identificado con DNI: 06813643, esta autorizado por nuestra empresa textil de la matta para que realice su trabajo de investigación Modelo predictivo del proceso de ventas utilizando inteligencia de negocios y data analytics para su tesis profesional.

Atte.


CENTRO TEXTIL DE LA MATTSA S.A.
MARISOL VERONICA MORA MEDINA
GERENTE GENERAL
MARISOL VERONICA MORA MEDINA
GERENTE GENERAL

Jr Sebastian barranca No 1441, La victoria - Lima

Anexo 3. Instrumentos de recolección de datos - Variable Independiente

Indicador - Tiempo Total del proceso

Promedio de Tiempo de respuesta	
Ítem	Valor
Velocidad	
Tiempo	
CPU	
Memoria	
Disco	

Nota. Fuente: elaboración propia.

Indicador - Memoria consumida durante un proceso

Consumo de Memoria	
Ítem	Valor
Uso	
Disponibilidad	
Confirmada	
En caché	
Tiempo	

Nota. Fuente: elaboración propia.

Indicador - CPU consumido durante un proceso

Consumo de CPU	
Ítem	Valor
Uso	
Velocidad	
Procesos	
Subprocesos	
Tiempo	

Nota. Fuente: elaboración propia.

Anexo 4. Instrumentos de recolección de datos - Variable Dependiente

Indicador - Error absoluto medio (MAE)

Épocas	Error Absoluto Medio del test (MAE)
	1
	2
	3
	4
	5
	6
	7
	8
	9
	10
	11
	12
	.
	.
	26
	27
	28
	29
Promedio	
Desviación Estándar	
Media	

Indicador - Error cuadrático medio (MSE)

Épocas	Error Cuadrático Medio del test (MSE)
	1
	2
	3
	4
	5
	6
	7
	8
	9
	10
	11
	12
	.
	.
	.
	27
	28
	29
Promedio	
Desviación Estándar	
Media	

Indicador - Error absoluto relativo (MAPE)

Épocas	Error Absoluto Relativo (MAPE)
	1
	2
	3
	4
	5
	6
	7
	8
	9
	10
	11
	12
	.
	.
	.
	27
	28
	29
Promedio	
Desviación Estándar	
Media	

Anexo 5. Instrumentos de recolección de datos previos

Guía de Entrevista N° 01

Entrevistado: Sr Jaime Matta	Fecha:
Entrevistador: Santiago Carreño G.	Área: Gerencia General
Objetivo: Analizar los Procesos de Ventas en los Puntos de Ventas La empresa Centro Textil de la Matta	
Dirigido a: Gerente General de la Empresa	
Preguntas:	
1. En qué medida le brinda información actualizada los puntos de ventas en la toma de decisiones.	
2. Defina Ud. Cuáles son los indicadores actuales que la empresa necesita como soporte para la toma de decisiones gerenciales de la empresa Centro Textil de la Matta y su impacto.	
3. Cómo calificaría Ud. Los nuevos resultados que brinda un nuevo sistema de gestión de procesos de ventas.	

Guía de Entrevista N° 02

Entrevistado: Sr Carlos Obando	Fecha:
Entrevistador: Santiago Carreño G.	Área: Almacén y Producción
Objetivo: Analizar los Procesos de Ventas en los Puntos de Ventas La empresa Centro Textil de la Matta	

Dirigido a: Administrador de Ventas y Almacenes
<p>Preguntas:</p> <p>1.En qué medida le brinda información actualizada los almacenes y de los puntos de ventas para proyectarse en la producción de nuevos productos</p> <p>2. Que indicadores actuales son realmente eficaces para la toma de decisiones en crear nuevos productos.</p> <p>3.Cómo calificaría Ud. Los nuevos resultados que brinda un nuevo sistema de gestión de procesos de ventas que considere la gestión de inventarios.</p>

Guía de Entrevista N° 03

Entrevistado: Ing. Luis Paredes Rosario	Fecha:
Entrevistador: Santiago Carreño G.	Área: Sistemas
Objetivo: Analizar los Procesos de información de data de ventas de la empresa Centro Textil de la Matta	
Dirigido a: Encargado de Sistemas	

- 1.Consultas y Referencias para la obtención y migración de información de datos
- 2.Tipo de Base de Datos y Lenguaje de Programación
- 3.Aprendizaje automático para pronostico

Anexo 6. Data de ventas de la empresa Centro Textil de la Matta SAC, 2017

1	A	B	C	D	E	F	G	H	I	J	K	L
	ANIO	MON	DOCU	FECHA	CTACTE	RAZON	CODIGO	DESCRIP	VV	CANT		
45329	2017	7	F001000005	3/07/2017	13484	INDUSTRIA TEXTIL DEL PACIFICO S.A.	T4110A10999	MINI JACQUARD		13.25	0.2660	
45331	2017	7	F001000007	3/07/2017	13484	INDUSTRIA TEXTIL DEL PACIFICO S.A.	TE600500999	CUELLOS BIRDSEYE		7.92	0.1010	
45333	2017	7	F001000007	3/07/2017	13484	INDUSTRIA TEXTIL DEL PACIFICO S.A.	TE600500999	CUELLOS BIRDSEYE		33.50	0.4270	
45336	2017	7	F001000007	3/07/2017	13484	INDUSTRIA TEXTIL DEL PACIFICO S.A.	T4143A49000	MINI JACQUARD MONDRIAN		213.09	4.6430	
45337	2017	7	F001000008	3/07/2017	13484	INDUSTRIA TEXTIL DEL PACIFICO S.A.	T1360A20000	SUPLEX SOLIDO HEAVY		3.34	0.1350	
45339	2017	7	F001000009	3/07/2017	13484	INDUSTRIA TEXTIL DEL PACIFICO S.A.	T1360A20000	SUPLEX SOLIDO HEAVY		3.34	0.1350	
45341	2017	7	F001000006	3/07/2017	13484	INDUSTRIA TEXTIL DEL PACIFICO S.A.	TE600500999	CUELLOS BIRDSEYE		33.50	0.4270	
45343	2017	7	F001000036	5/07/2017	13484	INDUSTRIA TEXTIL DEL PACIFICO S.A.	TE600500999	CUELLOS BIRDSEYE		0.00	0.0000	
45346	2017	7	F001000037	5/07/2017	13484	INDUSTRIA TEXTIL DEL PACIFICO S.A.	TE600500999	CUELLOS BIRDSEYE		0.00	0.0000	
45347	2017	7	F001000010	3/07/2017	33579	HILOS & ESTILOS S.A.C.	T4147A25999	DRY 1.60		50.92	1.1190	
45349	2017	7	F001000010	3/07/2017	33579	HILOS & ESTILOS S.A.C.	T4147A25999	DRY 1.60		50.87	1.1180	
45352	2017	7	F001000010	3/07/2017	33579	HILOS & ESTILOS S.A.C.	T4147A25999	DRY 1.60		51.01	1.1210	
45354	2017	7	F001000010	3/07/2017	33579	HILOS & ESTILOS S.A.C.	T4147A25999	DRY 1.60		51.42	1.1300	
45355	2017	7	F001000010	3/07/2017	33579	HILOS & ESTILOS S.A.C.	T4147A25999	DRY 1.60		51.01	1.1210	
45358	2017	7	F001000010	3/07/2017	33579	HILOS & ESTILOS S.A.C.	T4147A25999	DRY 1.60		50.19	1.1030	
45359	2017	7	F001000010	3/07/2017	33579	HILOS & ESTILOS S.A.C.	T4147A25999	DRY 1.60		50.47	1.1090	
45362	2017	7	F001000010	3/07/2017	33579	HILOS & ESTILOS S.A.C.	T4147A25999	DRY 1.60		50.47	1.1090	
45364	2017	7	F001000010	3/07/2017	33579	HILOS & ESTILOS S.A.C.	T4147A25999	DRY 1.60		50.87	1.1180	
45365	2017	7	F001000010	3/07/2017	33579	HILOS & ESTILOS S.A.C.	T4147A25999	DRY 1.60		50.37	1.1070	
45367	2017	7	F001000010	3/07/2017	33579	HILOS & ESTILOS S.A.C.	T4147A25999	DRY 1.60		50.83	1.1170	
45370	2017	7	F001000010	3/07/2017	33579	HILOS & ESTILOS S.A.C.	T4147A25999	DRY 1.60		47.87	1.0520	
45371	2017	7	F001000010	3/07/2017	33579	HILOS & ESTILOS S.A.C.	T4147A25999	DRY 1.60		50.78	1.1160	
45373	2017	7	F001000010	3/07/2017	33579	HILOS & ESTILOS S.A.C.	T4147A25999	DRY 1.60		50.19	1.1030	
45376	2017	7	F001000010	3/07/2017	33579	HILOS & ESTILOS S.A.C.	T4147A25999	DRY 1.60		49.92	1.0970	
45377	2017	7	F001000010	3/07/2017	33579	HILOS & ESTILOS S.A.C.	T4147A25999	DRY 1.60		51.01	1.1210	
45380	2017	7	F001000010	3/07/2017	33579	HILOS & ESTILOS S.A.C.	T4147A25999	DRY 1.60		51.19	1.1250	
45382	2017	7	F001000010	3/07/2017	33579	HILOS & ESTILOS S.A.C.	T4147A25999	DRY 1.60		51.15	1.1240	
45383	2017	7	F001000011	3/07/2017	20013	BLADE S.A.C	T4145A19999	DRY		46.97	0.9730	
45385	2017	7	F001000011	3/07/2017	20013	BLADE S.A.C	T4145A19999	DRY		47.89	0.9920	
45387	2017	7	F001000011	3/07/2017	20013	BLADE S.A.C	T4145A19999	DRY		45.28	0.9380	

Se encontraron 81576 de 386777 registros

Anexo 7. Data de ventas de la empresa Centro Textil de la Matta SAC, 2018

	A	B	C	D	E	F	G	H	I	J	K	L
1	ANIO	IV	DOCU	FECHA	CTACTE	RAZON	CODIGO	DESCRIP	VV	CANT		
88616	2018	1	F005007509	2/01/2018	70232	ETSAQC S.A.C	D0210900999	DESP. CARTON		6.86	80.9550	
88617	2018	1	F005007509	2/01/2018	70232	ETSAQC S.A.C	D0213100999	DESP. PLASTICO		0.72	2.1100	
88618	2018	1	F005007510	2/01/2018	6B082	ALANIZ JORGE PABLO	T1361A25999	TEXFLEX PESADO PEACH DO		1.53	0.0500	
88619	2018	1	F005007511	2/01/2018	44800	TITAN FASHION S.A.C.	T4145A19999	DRY		0.50	0.0320	
88620	2018	1	F005007511	2/01/2018	44800	TITAN FASHION S.A.C.	T4113A23999	MESH CARIBE		0.52	0.0280	
88621	2018	1	F005007511	2/01/2018	44800	TITAN FASHION S.A.C.	T5629A50999	WIN SPORT		0.31	0.0180	
88622	2018	1	F005007512	2/01/2018	69591	NOVABLUE SOURCING S.A.C.	T1552A15999	POLYFIX		4.14	0.2080	
88623	2018	1	F005007513	2/01/2018	69591	NOVABLUE SOURCING S.A.C.	T4147A25999	DRY 1.60		0.56	0.0400	
88624	2018	1	B005001540	2/01/2018	6B837	ALVAREZ ROMAN ROSA MELVA	T1356A19999	TEXFLEX PESADO NEW		19.82	0.9170	
88625	2018	1	B005001540	2/01/2018	6B837	ALVAREZ ROMAN ROSA MELVA	T1356A19999	TEXFLEX PESADO NEW		20.29	0.9390	
88626	2018	1	B005001541	2/01/2018	6B837	ALVAREZ ROMAN ROSA MELVA	T1361A28999	TEXFLEX PESADO CON ORILL		20.72	0.9980	
88627	2018	1	B005001541	2/01/2018	6B837	ALVAREZ ROMAN ROSA MELVA	T1361A28999	TEXFLEX PESADO CON ORILL		20.72	0.9980	
88628	2018	1	F005007514	2/01/2018	6B082	ALANIZ JORGE PABLO	T1361A28999	TEXFLEX PESADO CON ORILL		10.60	0.5210	
88629	2018	1	F005007515	2/01/2018	17409	IMPORTADORA EXPORTADORA SEBAST	T1301A23999	TEXFLEX LISTADO FEED		118.48	2.8840	
88630	2018	1	F005007516	2/01/2018	17409	IMPORTADORA EXPORTADORA SEBAST	T1301A23999	TEXFLEX LISTADO FEED		126.57	3.0810	
88631	2018	1	F005007517	2/01/2018	17409	IMPORTADORA EXPORTADORA SEBAST	T1301A23999	TEXFLEX LISTADO FEED		127.15	3.0950	
88632	2018	1	F005007518	2/01/2018	17409	IMPORTADORA EXPORTADORA SEBAST	T1301A23999	TEXFLEX LISTADO FEED		112.52	2.7390	
88633	2018	1	F005007519	2/01/2018	17409	IMPORTADORA EXPORTADORA SEBAST	T1301A23999	TEXFLEX LISTADO FEED		114.04	2.7760	
88634	2018	1	F005007522	2/01/2018	17409	IMPORTADORA EXPORTADORA SEBAST	T1301A23999	TEXFLEX LISTADO FEED		117.99	2.8720	
88635	2018	1	F005007523	2/01/2018	17409	IMPORTADORA EXPORTADORA SEBAST	T1301A23999	TEXFLEX LISTADO FEED		111.86	2.4980	
88636	2018	1	B005001542	2/01/2018	6A898	MONTALVO SOTO GUILLERMO	DT995001999	DESP. TELA RETAZO CAT.01		0.43	0.1700	
88637	2018	1	B005001542	2/01/2018	6A898	MONTALVO SOTO GUILLERMO	DT995002999	DESP. TELA RETAZO CAT.02		0.84	0.4950	
88638	2018	1	B005001543	2/01/2018	6B000	CAMPOS GUTIERREZ FORTUNITA	DT995001999	DESP. TELA RETAZO CAT.01		0.13	0.0750	
88639	2018	1	B005001543	2/01/2018	6B000	CAMPOS GUTIERREZ FORTUNITA	DT995002999	DESP. TELA RETAZO CAT.02		0.94	0.3700	
88640	2018	1	B005001543	2/01/2018	6B000	CAMPOS GUTIERREZ FORTUNITA	DT995003999	DESP. TELA RETAZO CAT.03		0.43	0.1275	
88641	2018	1	B005001544	2/01/2018	6D674	LOPEZ MONDOÑEDO ALEJANDRINA	DT995001999	DESP. TELA RETAZO CAT.01		0.24	0.0950	
88642	2018	1	B005001544	2/01/2018	6D674	LOPEZ MONDOÑEDO ALEJANDRINA	DT995002999	DESP. TELA RETAZO CAT.02		0.24	0.0925	
88643	2018	1	B005001544	2/01/2018	6D674	LOPEZ MONDOÑEDO ALEJANDRINA	DT995003999	DESP. TELA RETAZO CAT.03		0.06	0.0350	
88644	2018	1	F005007520	2/01/2018	13484	INDUSTRIA TEXTIL DEL PACIFICO S.A.	T1355A20999	JERSEY CATIONICO		13.80	0.4780	
88645	2018	1	F005007521	2/01/2018	25277	SOUTHERN TEXTILE NETWORK S A C	T3801A10999	TECH PLOUJE C/SPADEX		92.88	2.8060	

Se encontraron 90371 de 386777 registros

Anexo 8. Data de ventas de la empresa Centro Textil de la Matta SAC, 2019

A1	A	B	C	D	E	F	G	H	I	J	K	L
	ANIO	DOCU	FECHA	CTACTE	RAZON	CODIGO	DESCRIP	VV	CANT			
2	2019	1 F005022200	2/01/2019	70232	ETSAQC S.A.C	D0210900999	DESP. CARTON		1.48	17.4800		
3	2019	1 F005022200	2/01/2019	70232	ETSAQC S.A.C	D0213100999	DESP. PLASTICO		0.57	1.6800		
4	2019	1 F005022201	2/01/2019	31669	TRUJILLO GUZMAN VERONICA	T1301A43999	TEXFLEX XS ONE PEACH		29.41	1.0680		
5	2019	1 F005022201	2/01/2019	31669	TRUJILLO GUZMAN VERONICA	T1301A43999	TEXFLEX XS ONE PEACH		29.72	1.0790		
6	2019	1 F005022201	2/01/2019	31669	TRUJILLO GUZMAN VERONICA	T1301A43999	TEXFLEX XS ONE PEACH		29.41	1.0680		
7	2019	1 F005022202	2/01/2019	98291	DISTRIBUIDORA SUPER SPORT S.A.C	T4143A52999	MESH DRY NYLON MS		20.00	0.9390		
8	2019	1 F005022202	2/01/2019	98291	DISTRIBUIDORA SUPER SPORT S.A.C	T4143A52999	MESH DRY NYLON MS		24.01	1.1270		
9	2019	1 F005022202	2/01/2019	98291	DISTRIBUIDORA SUPER SPORT S.A.C	T4143A52999	MESH DRY NYLON MS		24.62	1.1560		
10	2019	1 F005022202	2/01/2019	98291	DISTRIBUIDORA SUPER SPORT S.A.C	T4143A52999	MESH DRY NYLON MS		23.90	1.1220		
11	2019	1 F005022202	2/01/2019	98291	DISTRIBUIDORA SUPER SPORT S.A.C	T4143A52999	MESH DRY NYLON MS		20.32	0.9540		
12	2019	1 F005022203	2/01/2019	98291	DISTRIBUIDORA SUPER SPORT S.A.C	T1301A44999	TEXFLEX PESADO XS NEW		20.00	1.0680		
13	2019	1 F005022204	2/01/2019	98291	DISTRIBUIDORA SUPER SPORT S.A.C	T1301A44999	TEXFLEX PESADO XS NEW		17.12	0.9140		
14	2019	1 F005022204	2/01/2019	98291	DISTRIBUIDORA SUPER SPORT S.A.C	T1301A44999	TEXFLEX PESADO XS NEW		17.27	0.9220		
15	2019	1 F005022204	2/01/2019	98291	DISTRIBUIDORA SUPER SPORT S.A.C	T1301A44999	TEXFLEX PESADO XS NEW		21.30	1.1370		
16	2019	1 F005022204	2/01/2019	98291	DISTRIBUIDORA SUPER SPORT S.A.C	T1301A44999	TEXFLEX PESADO XS NEW		20.21	1.0790		
17	2019	1 F005022204	2/01/2019	98291	DISTRIBUIDORA SUPER SPORT S.A.C	T1301A44999	TEXFLEX PESADO XS NEW		20.81	1.1110		
18	2019	1 F005022205	2/01/2019	6D745	CORPORACION JEDROK S.A.C.	T4113A23999	MESH CARIBE		16.61	0.9800		
19	2019	1 F005022205	2/01/2019	6D745	CORPORACION JEDROK S.A.C.	T4113A23999	MESH CARIBE		16.61	0.9800		
20	2019	1 F005022205	2/01/2019	6D745	CORPORACION JEDROK S.A.C.	T4113A23999	MESH CARIBE		16.63	0.9810		
21	2019	1 F005022206	2/01/2019	34091	MAMANI COILA NANCY	T0143A15999	AMAR ANTILOPE		17.62	1.0770		
22	2019	1 F005022206	2/01/2019	34091	MAMANI COILA NANCY	T0143A15999	AMAR ANTILOPE		18.41	1.1250		
23	2019	1 F005022206	2/01/2019	34091	MAMANI COILA NANCY	T0143A15999	AMAR ANTILOPE		17.51	1.0700		
24	2019	1 F005022206	2/01/2019	34091	MAMANI COILA NANCY	T0143A15999	AMAR ANTILOPE		17.62	1.0770		
25	2019	1 F005022207	2/01/2019	34091	MAMANI COILA NANCY	T0143A15999	AMAR ANTILOPE		16.80	1.0270		
26	2019	1 F005022207	2/01/2019	34091	MAMANI COILA NANCY	T0143A15999	AMAR ANTILOPE		19.60	1.1980		
27	2019	1 F005022207	2/01/2019	34091	MAMANI COILA NANCY	T0143A15999	AMAR ANTILOPE		18.32	1.1200		
28	2019	1 F005022207	2/01/2019	34091	MAMANI COILA NANCY	T0143A15999	AMAR ANTILOPE		18.41	1.1250		
29	2019	1 F005022207	2/01/2019	34091	MAMANI COILA NANCY	T0143A15999	AMAR ANTILOPE		16.46	1.0060		
30	2019	1 F005022207	2/01/2019	34091	MAMANI COILA NANCY	T0143A15999	AMAR ANTILOPE		18.85	1.1520		
31	2019	1 F005022207	2/01/2019	34091	MAMANI COILA NANCY	T0143A15999	AMAR ANTILOPE		18.27	1.1170		

Se encontraron 88407 de 386777 registros

Anexo 9. Data de ventas de la empresa Centro Textil de la Matta SAC, 2020

1	ANIO	M	DOCU	FECHA	CTACTE	RAZON	CODIGO	DESCRIP	VV	CANT		
260356	2020	1	F005035890	6/01/2020	98410	ORELLANA PASCUAL ELOY FREDY	T1301A23999	TEXFLEX LISTADO FEED		3.15	0.0680	
260357	2020	1	B005007710	6/01/2020	6A550	ARAPA APAZA DIULANA	T0143A15999	AMAR ANTILOPE		19.55	1.1650	
260358	2020	1	B005007710	6/01/2020	6A550	ARAPA APAZA DIULANA	T0143A15999	AMAR ANTILOPE		19.31	1.1510	
260359	2020	1	B005007710	6/01/2020	6A550	ARAPA APAZA DIULANA	T0143A15999	AMAR ANTILOPE		18.68	1.1130	
260360	2020	1	B005007710	6/01/2020	6A550	ARAPA APAZA DIULANA	T0143A15999	AMAR ANTILOPE		18.42	1.0980	
260361	2020	1	B005007710	6/01/2020	6A550	ARAPA APAZA DIULANA	T0143A15999	AMAR ANTILOPE		18.66	1.1120	
260362	2020	1	B005007710	6/01/2020	6A550	ARAPA APAZA DIULANA	T0143A15999	AMAR ANTILOPE		18.69	1.1140	
260363	2020	1	B005007710	6/01/2020	6A550	ARAPA APAZA DIULANA	T0143A15999	AMAR ANTILOPE		18.48	1.1010	
260364	2020	1	B005007710	6/01/2020	6A550	ARAPA APAZA DIULANA	T0143A15999	AMAR ANTILOPE		18.91	1.1270	
260365	2020	1	B005007710	6/01/2020	6A550	ARAPA APAZA DIULANA	T0143A15999	AMAR ANTILOPE		19.25	1.1470	
260366	2020	1	B005007710	6/01/2020	6A550	ARAPA APAZA DIULANA	T0143A15999	AMAR ANTILOPE		19.01	1.1330	
260367	2020	1	B005007710	6/01/2020	6A550	ARAPA APAZA DIULANA	T0143A15999	AMAR ANTILOPE		18.83	1.1220	
260368	2020	1	B005007710	6/01/2020	6A550	ARAPA APAZA DIULANA	T0143A15999	AMAR ANTILOPE		18.58	1.1070	
260369	2020	1	F005035891	6/01/2020	6A779	ARIAS QUISPE ANA ESMERALDA	T0143A15999	AMAR ANTILOPE		19.23	1.1290	
260370	2020	1	F005035891	6/01/2020	6A779	ARIAS QUISPE ANA ESMERALDA	T0143A15999	AMAR ANTILOPE		19.02	1.1170	
260371	2020	1	F005035891	6/01/2020	6A779	ARIAS QUISPE ANA ESMERALDA	T0143A15999	AMAR ANTILOPE		18.85	1.1070	
260372	2020	1	F005035891	6/01/2020	6A779	ARIAS QUISPE ANA ESMERALDA	T0143A15999	AMAR ANTILOPE		18.90	1.1100	
260373	2020	1	F005035891	6/01/2020	6A779	ARIAS QUISPE ANA ESMERALDA	T0143A15999	AMAR ANTILOPE		19.24	1.1300	
260374	2020	1	F005035891	6/01/2020	6A779	ARIAS QUISPE ANA ESMERALDA	T0143A15999	AMAR ANTILOPE		19.33	1.1350	
260375	2020	1	F005035891	6/01/2020	6A779	ARIAS QUISPE ANA ESMERALDA	T0143A15999	AMAR ANTILOPE		19.12	1.1230	
260376	2020	1	F005035891	6/01/2020	6A779	ARIAS QUISPE ANA ESMERALDA	T0143A15999	AMAR ANTILOPE		18.63	1.0940	
260377	2020	1	F005035891	6/01/2020	6A779	ARIAS QUISPE ANA ESMERALDA	T0143A15999	AMAR ANTILOPE		18.84	1.1060	
260378	2020	1	F005035891	6/01/2020	6A779	ARIAS QUISPE ANA ESMERALDA	T0143A15999	AMAR ANTILOPE		18.90	1.1100	
260379	2020	1	F005035891	6/01/2020	6A779	ARIAS QUISPE ANA ESMERALDA	T0143A15999	AMAR ANTILOPE		18.85	1.1070	
260380	2020	1	F005035891	6/01/2020	6A779	ARIAS QUISPE ANA ESMERALDA	T0143A15999	AMAR ANTILOPE		18.89	1.1090	
260381	2020	1	F005035891	6/01/2020	6A779	ARIAS QUISPE ANA ESMERALDA	T0143A15999	AMAR ANTILOPE		18.82	1.1050	
260382	2020	1	F005035892	6/01/2020	6D873	J.E.M. TEXTILES E.I.R.L	T1552A15999	POLYFIX		22.87	1.1390	
260383	2020	1	F005035892	6/01/2020	6D873	J.E.M. TEXTILES E.I.R.L	T1552A15999	POLYFIX		22.79	1.1350	
260384	2020	1	F005035892	6/01/2020	6D873	J.E.M. TEXTILES E.I.R.L	T1552A15999	POLYFIX		23.13	1.1520	
260385	2020	1	F005035892	6/01/2020	6D873	J.E.M. TEXTILES E.I.R.L	T1552A15999	POLYFIX		21.77	1.0840	

Se encontraron 68755 de 386777 registros

Anexo 10. Data de ventas de la empresa Centro Textil de la Matta SAC, 2021

1	A	B	C	D	E	F	G	H	I	J	K	L
	ANIO	M	DOCU	FECHA	CTACTE	RAZON	CODIGO	DESCRIP	VV	CANT		
329111	2021	3	FF00500047713	#####	20101600735	ALMERIZ S.A.	T0103A17000	OTTOMAN CATIONICO		38.71	0.9550	
329112	2021	3	FF00500047713	#####	20101600735	ALMERIZ S.A.	T0103A17000	OTTOMAN CATIONICO		40.21	0.9920	
329113	2021	3	FF00500047713	#####	20101600735	ALMERIZ S.A.	T0103A17000	OTTOMAN CATIONICO		39.89	0.9840	
329114	2021	3	FF00500047713	#####	20101600735	ALMERIZ S.A.	T0103A17000	OTTOMAN CATIONICO		39.61	0.9770	
329115	2021	3	FF00500047713	#####	20101600735	ALMERIZ S.A.	T0103A17000	OTTOMAN CATIONICO		39.73	0.9800	
329116	2021	3	FF00500047713	#####	20101600735	ALMERIZ S.A.	T0103A17000	OTTOMAN CATIONICO		40.25	0.9930	
329117	2021	3	FF00500047713	#####	20101600735	ALMERIZ S.A.	T0103A17000	OTTOMAN CATIONICO		39.81	0.9820	
329118	2021	3	FF00500047713	#####	20101600735	ALMERIZ S.A.	T0103A17000	OTTOMAN CATIONICO		41.87	1.0330	
329119	2021	3	FF00500047713	#####	20101600735	ALMERIZ S.A.	T0103A17000	OTTOMAN CATIONICO		39.77	0.9810	
329120	2021	3	FF00500047713	#####	20101600735	ALMERIZ S.A.	T0103A17000	OTTOMAN CATIONICO		41.19	1.0160	
329121	2021	3	FF00500047713	#####	20101600735	ALMERIZ S.A.	T0103A17000	OTTOMAN CATIONICO		39.77	0.9810	
329122	2021	3	FF00500047713	#####	20101600735	ALMERIZ S.A.	T0103A17000	OTTOMAN CATIONICO		39.93	0.9850	
329123	2021	3	FF00500047843	#####	20376729126	SOUTHERN TEXTILE NETWORK S.A.C.	T1356A13999	TEXFLEX CATIONICO		36.48	0.7790	
329124	2021	3	FF00500047843	#####	20376729126	SOUTHERN TEXTILE NETWORK S.A.C.	T1356A13999	TEXFLEX CATIONICO		44.15	0.9430	
329125	2021	4	FF00500048012	5/04/2021	20392817167	TANDEM TEXTIL S.A.C.	T1301A23999	TEXFLEX LISTADO FEED		75.85	1.0350	
329126	2021	3	FF002008511	#####	20112316249	INDUSTRIA TEXTIL DEL PACIFICO S.A.	T4143A25999	MESH LLANO		374.10	12.4700	
329127	2021	3	FF00500047846	#####	20522268489	VERANA S.A.C.	T2021A05999	PLUSH LLANO		41.17	0.7050	
329128	2021	3	FF00500047846	#####	20522268489	VERANA S.A.C.	T2021A05999	PLUSH LLANO		34.11	0.5840	
329129	2021	3	FF00500047846	#####	20522268489	VERANA S.A.C.	T2021A05999	PLUSH LLANO		30.73	0.5262	
329130	2021	4	FF002008709	8/04/2021	20381654614	IMPORTADORA EXPORTADORA SEBAST	T1301A23999	TEXFLEX LISTADO FEED		241.30	5.6700	
329131	2021	4	FF002008709	8/04/2021	20381654614	IMPORTADORA EXPORTADORA SEBAST	T1301A23999	TEXFLEX LISTADO FEED		239.60	5.6300	
329132	2021	4	FF002008710	8/04/2021	20112316249	INDUSTRIA TEXTIL DEL PACIFICO S.A.	T4113A12999	MESH ALPHA		22.15	0.9400	
329133	2021	4	FF002008711	8/04/2021	20112316249	INDUSTRIA TEXTIL DEL PACIFICO S.A.	T4113A12999	MESH ALPHA		22.62	0.9600	
329134	2021	4	FF002008711	8/04/2021	20112316249	INDUSTRIA TEXTIL DEL PACIFICO S.A.	T4113A12999	MESH ALPHA		296.18	12.5700	
329135	2021	3	FF00500047884	#####	10427702681	HUAMANI HUILLCA YUDITH	T1552A15999	POLYFIX		23.22	1.1320	
329136	2021	3	FF00500047884	#####	10427702681	HUAMANI HUILLCA YUDITH	T1552A15999	POLYFIX		23.50	1.1460	
329137	2021	3	FF00500047884	#####	10427702681	HUAMANI HUILLCA YUDITH	T1552A15999	POLYFIX		22.68	1.1060	
329138	2021	3	FF00500047884	#####	10427702681	HUAMANI HUILLCA YUDITH	T1552A15999	POLYFIX		23.24	1.1330	
329139	2021	3	FF00500047884	#####	10427702681	HUAMANI HUILLCA YUDITH	T1552A15999	POLYFIX		23.44	1.1430	
329140	2021	3	FF00500047884	#####	10427702681	HUAMANI HUILLCA YUDITH	T1552A15999	POLYFIX		22.64	1.1040	

Se encontraron 57668 de 386777 registros

Anexo 11. Resultados del indicador “Tiempo total del proceso”

```
Epoch 1/40
170/170 [=====] - 3s 8ms/step -
0.0613 - val_mae: 0.1987 - val_mape: 60.5367
Epoch 2/40
170/170 [=====] - 1s 3ms/step -
0.0561 - val_mae: 0.1858 - val_mape: 57.4295
Epoch 3/40
170/170 [=====] - 1s 4ms/step -
0.0542 - val_mae: 0.1815 - val_mape: 56.4703
Epoch 4/40
170/170 [=====] - 1s 3ms/step -
0.0532 - val_mae: 0.1804 - val_mape: 56.5004
Epoch 5/40
170/170 [=====] - 1s 3ms/step -
0.0533 - val_mae: 0.1828 - val_mape: 57.5559
Epoch 6/40
170/170 [=====] - 1s 3ms/step -
0.0512 - val_mae: 0.1792 - val_mape: 55.1118
Epoch 7/40
170/170 [=====] - 1s 3ms/step -
0.0530 - val_mae: 0.1844 - val_mape: 57.7787
Epoch 8/40
170/170 [=====] - 1s 4ms/step -
0.0525 - val_mae: 0.1850 - val_mape: 57.3134
```

```
Epoch 9/40
170/170 [=====] - 1s 3ms/step
0.0554 - val_mae: 0.1920 - val_mape: 60.7164
Epoch 10/40
170/170 [=====] - 1s 3ms/step
0.0509 - val_mae: 0.1821 - val_mape: 55.1257
Epoch 11/40
170/170 [=====] - 1s 4ms/step
0.0507 - val_mae: 0.1817 - val_mape: 54.6322
Epoch 12/40
170/170 [=====] - 1s 3ms/step
0.0555 - val_mae: 0.1934 - val_mape: 60.8229
Epoch 13/40
170/170 [=====] - 1s 4ms/step
0.0560 - val_mae: 0.1950 - val_mape: 61.2007
Epoch 14/40
170/170 [=====] - 1s 4ms/step
0.0552 - val_mae: 0.1932 - val_mape: 60.2302
Epoch 15/40
170/170 [=====] - 1s 3ms/step
0.0527 - val_mae: 0.1873 - val_mape: 57.0958
Epoch 16/40
```

```
Epoch 16/40
170/170 [=====] - 1s 4ms/step - loss: 0.1707 - mse: 0.0534 - mae: 0.1707 - mape: 62.9524 - val_loss: 0.1875 - val_mse:
0.0527 - val_mae: 0.1875 - val_mape: 57.2655
Epoch 17/40
170/170 [=====] - 1s 4ms/step - loss: 0.1706 - mse: 0.0536 - mae: 0.1706 - mape: 63.3176 - val_loss: 0.1873 - val_mse:
0.0526 - val_mae: 0.1873 - val_mape: 56.9654
Epoch 18/40
170/170 [=====] - 1s 4ms/step - loss: 0.1710 - mse: 0.0536 - mae: 0.1710 - mape: 63.6376 - val_loss: 0.1869 - val_mse:
0.0524 - val_mae: 0.1869 - val_mape: 56.3483
Epoch 19/40
170/170 [=====] - 1s 3ms/step - loss: 0.1704 - mse: 0.0531 - mae: 0.1704 - mape: 62.8166 - val_loss: 0.1879 - val_mse:
0.0528 - val_mae: 0.1879 - val_mape: 57.0710
Epoch 20/40
170/170 [=====] - 1s 3ms/step - loss: 0.1711 - mse: 0.0535 - mae: 0.1711 - mape: 62.9738 - val_loss: 0.1977 - val_mse:
0.0566 - val_mae: 0.1977 - val_mape: 61.6268
Epoch 21/40
170/170 [=====] - 1s 4ms/step - loss: 0.1705 - mse: 0.0535 - mae: 0.1705 - mape: 63.1360 - val_loss: 0.1875 - val_mse:
0.0525 - val_mae: 0.1875 - val_mape: 56.3439
Epoch 22/40
170/170 [=====] - 1s 4ms/step - loss: 0.1710 - mse: 0.0536 - mae: 0.1710 - mape: 63.8251 - val_loss: 0.1918 - val_mse:
0.0543 - val_mae: 0.1918 - val_mape: 59.0035
```

```
Epoch 23/40
170/170 [=====] - 1s 4ms/step - loss: 0.1707 - mse: 0.0530 - mae: 0.1707 - mape: 62.9479 - val_loss: 0.1855 - val_mse:
0.0521 - val_mae: 0.1855 - val_mape: 55.6296
Epoch 24/40
170/170 [=====] - 1s 4ms/step - loss: 0.1700 - mse: 0.0528 - mae: 0.1700 - mape: 62.5102 - val_loss: 0.1965 - val_mse:
0.0562 - val_mae: 0.1965 - val_mape: 61.0891
Epoch 25/40
170/170 [=====] - 1s 4ms/step - loss: 0.1708 - mse: 0.0536 - mae: 0.1708 - mape: 63.5434 - val_loss: 0.2004 - val_mse:
0.0576 - val_mae: 0.2004 - val_mape: 62.5572
Epoch 26/40
170/170 [=====] - 1s 4ms/step - loss: 0.1701 - mse: 0.0532 - mae: 0.1701 - mape: 63.3296 - val_loss: 0.1822 - val_mse:
0.0512 - val_mae: 0.1822 - val_mape: 52.8926
Epoch 27/40
170/170 [=====] - 1s 3ms/step - loss: 0.1710 - mse: 0.0532 - mae: 0.1710 - mape: 62.4444 - val_loss: 0.1921 - val_mse:
0.0544 - val_mae: 0.1921 - val_mape: 58.7146
Epoch 28/40
170/170 [=====] - 1s 4ms/step - loss: 0.1707 - mse: 0.0530 - mae: 0.1707 - mape: 62.1483 - val_loss: 0.1995 - val_mse:
0.0574 - val_mae: 0.1995 - val_mape: 62.3083
Epoch 29/40
170/170 [=====] - 1s 3ms/step - loss: 0.1704 - mse: 0.0534 - mae: 0.1704 - mape: 63.7451 - val_loss: 0.1914 - val_mse:
0.0542 - val_mae: 0.1914 - val_mape: 58.4293
Epoch 30/40
170/170 [=====] - 1s 4ms/step - loss: 0.1699 - mse: 0.0530 - mae: 0.1699 - mape: 62.8347 - val_loss: 0.1877 - val_mse:
0.0525 - val_mae: 0.1877 - val_mape: 56.4977
```

```
Epoch 31/40
170/170 [=====] - 1s 3ms/step - loss: 0.1706 - mse: 0.0531 - mae: 0.1706 - mape: 62.9487 - val_loss: 0.1934 - val_mse:
0.0547 - val_mae: 0.1934 - val_mape: 59.2180
Epoch 32/40
170/170 [=====] - 1s 3ms/step - loss: 0.1702 - mse: 0.0530 - mae: 0.1702 - mape: 62.6729 - val_loss: 0.1934 - val_mse:
0.0548 - val_mae: 0.1934 - val_mape: 59.3098
Epoch 33/40
170/170 [=====] - 1s 3ms/step - loss: 0.1697 - mse: 0.0529 - mae: 0.1697 - mape: 62.2967 - val_loss: 0.1941 - val_mse:
0.0553 - val_mae: 0.1941 - val_mape: 59.9798
Epoch 34/40
170/170 [=====] - 1s 4ms/step - loss: 0.1699 - mse: 0.0530 - mae: 0.1699 - mape: 63.3047 - val_loss: 0.2003 - val_mse:
0.0578 - val_mae: 0.2003 - val_mape: 62.5526
Epoch 35/40
170/170 [=====] - 1s 3ms/step - loss: 0.1701 - mse: 0.0532 - mae: 0.1701 - mape: 63.4078 - val_loss: 0.1861 - val_mse:
0.0523 - val_mae: 0.1861 - val_mape: 55.2321
Epoch 36/40
170/170 [=====] - 1s 3ms/step - loss: 0.1696 - mse: 0.0526 - mae: 0.1696 - mape: 62.1931 - val_loss: 0.1890 - val_mse:
0.0532 - val_mae: 0.1890 - val_mape: 57.0384
Epoch 37/40
170/170 [=====] - 1s 3ms/step - loss: 0.1704 - mse: 0.0531 - mae: 0.1704 - mape: 63.0245 - val_loss: 0.1848 - val_mse:
0.0519 - val_mae: 0.1848 - val_mape: 54.3737
Epoch 38/40
170/170 [=====] - 1s 3ms/step - loss: 0.1700 - mse: 0.0531 - mae: 0.1700 - mape: 62.3802 - val_loss: 0.1865 - val_mse:
0.0525 - val_mae: 0.1865 - val_mape: 55.2400
```

```
Epoch 38/40
170/170 [=====] - 1s 3ms/step - loss: 0.1700 - mse: 0.0531 - mae: 0.1700 - mape: 62.3802 - val_loss: 0.1865 - val_mse:
0.0525 - val_mae: 0.1865 - val_mape: 55.2400
Epoch 39/40
170/170 [=====] - 1s 4ms/step - loss: 0.1700 - mse: 0.0530 - mae: 0.1700 - mape: 62.8092 - val_loss: 0.1968 - val_mse:
0.0563 - val_mae: 0.1968 - val_mape: 60.6661
Epoch 40/40
170/170 [=====] - 1s 3ms/step - loss: 0.1696 - mse: 0.0530 - mae: 0.1696 - mape: 63.1228 - val_loss: 0.1911 - val_mse:
0.0541 - val_mae: 0.1911 - val_mape: 57.6542
```


Anexo 12. Resultados del indicador “Memoria consumida durante un proceso”

Uso de Memoria al inicio del Entrenamiento

```
La cantidad de memoria usada actualmente es 0.0MB; el pico fue de 0.0MB
```

Código Python

```
# Aquí se activa el rastreo de memoria para saber el consumo de memoria en el modelo (antes de correr el modelo)
tracemalloc.start()
current, peak = tracemalloc.get_traced_memory()
print(f"La cantidad de memoria usada actualmente es {current / 10**6}MB; el pico fue de {peak / 10**6}MB")
tracemalloc.stop()
```

Memoria al término del Entrenamiento

```
La cantidad de memoria usada actualmente es 5.959833MB; el pico fue de 5.976187MB
```

Uso de Memoria al término del Entrenamiento

Finalizar el uso de Memoria

Código Python

```
print(x_train.shape, y_train.shape, x_val.shape, y_val.shape)
"""Se ejecutará nuestro modelo y mostraremos como se comporta la maquina despues de 40 épocas."""
# aquí se activa el rastreo de memoria para saber el consumo de memoria en el modelo (Cuando empieza a correr el modelo)
model = crear_modeloFF(pasos)

tracemalloc.start()
current, peak = tracemalloc.get_traced_memory()
t_0 = time.time()
history = model.fit(x_train, y_train, epochs=epochs, validation_data=(x_val, y_val), batch_size=pasos)
t_1 = time.time()

"""Se finaliza el proceso de uso de memoria en el entrenamiento"""
current, peak = tracemalloc.get_traced_memory()

print(f"La cantidad de memoria usada actualmente es {current / 10**6}MB; el pico fue de {peak / 10**6}MB")
tracemalloc.stop()
```

Anexo 13. Resultados del indicador “CPU consumido durante un proceso”

Uso de CPU al Inicio del Entrenamiento

```
CPU usado es: 34.2
```

Código Python

```
# Se activa el INICIO DE USO CPU  
print('CPU usado es: ', psutil.cpu_percent(4))
```

Uso de CPU al término del Entrenamiento

```
CPU usado es: 33.6
```

Código Python

```
... """se finaliza el uso de CPU en el entrenamiento"""  
... print("Métricas de entrenamiento:")  
... print('CPU usado es: ', psutil.cpu_percent(4))
```

```
Console 1/A x  
DataFrame INFO:  
<class 'pandas.core.frame.DataFrame'>  
DatetimeIndex: 1460 entries, 2017-01-02 to 2021-09-23  
Data columns (total 1 columns):  
#   Column  Non-Null Count  Dtype  
---  ---      -  
0   VV      1460 non-null    float64  
dtypes: float64(1)  
memory usage: 22.8 KB  
None  
  
La cantidad de memoria usada actualmente es 0.0MB; el pico fue de 0.0MB  
CPU usado es: 32.5  
(1414, 1, 7) (1414,) (30, 1, 7) (30,)  
Model: "sequential"
```

```
202/202 [=====] - 1s 3ms/step - loss: 0.1770 - mse: 0.0571 - mae: 0.1770 - mape: 82.1494 - val_loss: 0.1801 - val_mse: 0.0550 -  
val_mae: 0.1861 - val_mape: 220.5919  
Epoch 38/40  
202/202 [=====] - 1s 4ms/step - loss: 0.1780 - mse: 0.0571 - mae: 0.1780 - mape: 81.9219 - val_loss: 0.1978 - val_mse: 0.0559 -  
val_mae: 0.1978 - val_mape: 237.4084  
Epoch 39/40  
202/202 [=====] - 1s 3ms/step - loss: 0.1782 - mse: 0.0573 - mae: 0.1782 - mape: 82.1863 - val_loss: 0.1879 - val_mse: 0.0518 -  
val_mae: 0.1879 - val_mape: 223.6479  
Epoch 40/40  
202/202 [=====] - 1s 3ms/step - loss: 0.1776 - mse: 0.0569 - mae: 0.1776 - mape: 81.3446 - val_loss: 0.1949 - val_mse: 0.0546 -  
val_mae: 0.1949 - val_mape: 232.1928  
Métricas de entrenamiento:  
CPU usado es: 39.1  
La cantidad de memoria usada actualmente es 5.730132MB; el pico fue de 5.746438MB  
1/1 [=====] - 0s 105ms/step  
30
```

Anexo 14. Resultados de entrevistas a personal de Centro Textil de la Matta

Entrevista a Gerente General



Entrevistado: Sr Jaime Matta Entrevistador: Santiago Carreño G.	Fecha: 9/01/2021 Área: Gerencia General
Objetivo: Analizar los Procesos de Ventas en los Puntos de Ventas La empresa Centro Textil de la Matta	
Dirigido a: Gerente General de la Empresa	
Preguntas: 1. En qué medida le brinda información actualizada los puntos de ventas en la toma de decisiones. - Al tener actualizada toda la información en un sistema apoyará a tomar decisiones. - El sistema de puntos de ventas nos brindará resultados; datos que me permitan fabricar los productos más vendidos. - Podemos proyectar las compras de materia prima para fabricar. 2. Defina Ud. Cuáles son los indicadores actuales que la empresa necesita como soporte para la toma de decisiones gerenciales de la empresa Centro Textil de la Matta y su impacto. - Pronosticar las ventas de forma mensual. - Verificar de forma mensual el crecimiento de ventas. - El sistema actual permite apoyar la parte financiera. - El sistema nos brinda reportes para realizar el forecast de ventas de productos. 3. Cómo calificaría Ud. Los nuevos resultados que brinda un nuevo sistema de gestión de procesos de ventas. - La atención a nuestros clientes es rápida y eficiente. - Si nos brinda exactamente el monto total de ingresos de ventas; se realiza diariamente y también mensual. - Los documentos como facturas, boletas y guías son digitales y exactas.	


Firma

Entrevista a Administrador de Ventas y Almacenes


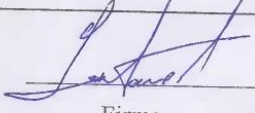


Entrevistado: Sr Carlos Obando Entrevistador: Santiago Carreño G.	Fecha: 6/01/2021 Área: Ventas y Almacenes
Objetivo: Analizar los Procesos de Ventas en los Puntos de Ventas La empresa Centro Textil de la Matta	
Dirigido a: Administrador de ventas y Almacenes	
Preguntas: 1. En qué medida le brinda información actualizada los almacenes y de los puntos de ventas para proyectarse en la producción de nuevos productos. <ul style="list-style-type: none">- En el sistema manual no se tenía la información actualizada y no había orden documentario.- En el sistema actual mecanizado, la información es fluida y actualizada.- Los reportes se obtienen de manera rápida y los resultados son exactos.	
2. Que indicadores actuales son realmente eficaces para la toma de decisiones en crear nuevos productos. <ul style="list-style-type: none">- Verificar el total de ventas diarias.- Proyectar las ventas de los productos.- Considerar los productos que más se venden.- Proyectar la fabricación de productos.	
3. Cómo calificaría Ud. Los nuevos resultados que brinda un nuevo sistema de gestión de procesos de ventas que considere la gestión de inventarios. <ul style="list-style-type: none">- Los resultados son exactos en el nuevo sistema.- El proceso de atención es rápido.- El cliente está más contento por la atención.	

O. Obando

Firma

Entrevista a Encargado de Sistemas

 Centro Textil de la Matta	
Entrevistado: Ing Luis Paredes R. Entrevistador: Santiago Carreño G.	Fecha: 13/01/2021 Área: Gerencia General
Objetivo: Analizar los Procesos de Ventas en los Puntos de Ventas La empresa Centro Textil de la Matta	
Dirigido a: Encargado de Sistemas	
Preguntas:	
1. Consultas y Referencias para la obtención y migración de información de datos. <i>Para poder migrar la información de ventas se ha considerado los años 2017, 2018, 2019, 2020 y parte del año 2021, migrando de SQL server y luego migrar a Hoja Excel.</i>	
2.-Tipo de Base de Datos y Lenguaje de Programación. <i>La Base de datos esta en SQL SERVER, 2014 EXPRESS EL sistema de ventas se encuentra diseñado en lenguaje de programación visual FOX VER 9.0.</i>	
3. Aprendizaje automático para pronostico <i>Es necesario contar con proyecciones de ventas futuras Apoyada merito a la toma de decisiones para la producción de los productos. y proyección de ingresos financieros y gerencia.</i>	
 Firma	

Anexo 15. Evidencias fotográficas.



Foto de entrevista con Ing. Luis Paredes Rosario, Encargado de Sistemas