Modelo dinámico para el pronóstico de la demanda en empresa de alimentos

Dynamic model for demand forecasting in a food company

Normeviz Depablo Vizcaya, Paola Martínez Valero, Francisco Figueredo Lugo

Palabras clave: redes neuronales, métodos de pronóstico, sistema de empuje, cadena de suministro
Key words: neural networks, forecasting methods, push system, supply chain

RESUMEN

El presente trabajo tiene como objetivo diseñar un modelo dinámico para el pronóstico de la demanda en una empresa productora de alimentos, con el fin de mejorar la precisión del proceso actual y disminuir la sobreestimación y subestimación de la demanda. Esta demanda genera importantes problemas para la empresa como: exceso de producción, ruptura de stock, insatisfacción de clientes, costos adicionales, entre otros. La metodología de la investigación es Proyecto Factible, Descriptivo, diseño de Campo y enfocado en una unidad de análisis. Se recopilaron datos históricos de ventas, se segmentaron los productos por el principio de Pareto y se diseñó una red neuronal recurrente multicapa con celdas LSTM combinada con métodos estadísticos tradicionales (Bootstrap). Se ajustaron los hiper parámetros del modelo y del entrenamiento. Se comparó el método de pronóstico actual de la empresa con el propuesto, realizando una evaluación estadística de los errores. El método propuesto obtuvo una precisión promedio del 94,11 % frente al 30,22 % del método actual, lo que representa una mejora promedio del 63,89 %.

ABSTRACT

The objective of this work is to design a dynamic model for demand forecasting in a food producing company, in order to improve the precision of the current process and reduce the overestimation and underestimation of demand. This generates substantial problems for the company such as: excess production, stock out, customer dissatisfaction and additional costs, among others. The research methodology is Feasible, Descriptive Project, field design and focused on a unit of analysis. Historical sales data was collected, products were segmented by the Pareto principle and a multilayer recurrent neural network was designed with LSTM cells combined with traditional statistical methods (Bootstrap). The model and training hyperparameters were adjusted. The company's current forecasting method was compared with the proposed one, performing a statistical evaluation of the errors. The proposed method obtained an average accuracy of 94.11 % compared to 30.22 % for the current method, which represents an average improvement of ki 63.89 %.

INTRODUCCIÓN

En el panorama empresarial actual, altamente competitivo y en constante evolución, el éxito de las empresas depende en gran medida de la toma de decisiones estratégicas en la gestión de la cadena de suministro. Como se citó en Chase y Jacobs (2009), Gates (1999) afirma que la supervivencia de la mayor parte de las empresas depende de las decisiones inteligentes sobre la cadena de suministro en especial de aplicar la tecnología e inteligencia para mejorar el desempeño de la misma.

En un mercado cambiante, las empresas deben pronosticar la demanda de productos, lo que puede resultar difícil, pero es esencial para evitar una reducción en los ingresos. Por lo tanto, la cadena de suministro debe ser ágil para responder rápidamente a los cambios en la demanda (Del Castillo, 2023). El autor resalta dos aspectos clave para que las empresas sean competitivas en un entorno cambiante: el pronóstico preciso de la demanda y la agilidad de la cadena de suministro.

Pronosticar es definido como el arte y la ciencia para predecir eventos futuros. Puede implicar el uso de datos históricos y su proyección hacia el futuro mediante algún modelo matemático. Puede ser una predicción subjetiva o intuitiva, o puede ser una combinación de ambos, es decir, un modelo matemático ajustado por el buen juicio del administrador (Render et al, 2007).

En el ámbito de la predicción de la demanda, el aprendizaje automático surge como una herramienta potente para aprender patrones no sólo lineales como los métodos tradicionales para pronóstico. De acuerdo con Sandoval (2018) el aprendizaje automático se encarga de generar algoritmos que tienen la capacidad de aprender y no tener que programarlos de manera explícita.

Una de las aplicaciones más poderosas y de mayor crecimiento de la inteligencia artificial es el aprendizaje profundo (en inglés, *deep learning*). Se trata de un subcampo del aprendizaje automático que se utiliza para resolver problemas muy complejos y que normalmente implican grandes cantidades de datos (Rouhiainen, 2018). En este sentido Arana, (2021) explica que las Redes Neuronales Recurrentes (RNNs) son una clase de aprendizaje profundo basada en los trabajos de David Rumelhart en 1986. Las RNN son conocidas por su capacidad para procesar y obtener información de datos secuenciales. Según Nacelle (2009) el flujo de activación de las RNN va en una sola dirección: hacia delante y sólo se procede a la propagación hacia atrás para adecuar los pesos de cada neurona.

Las RNN tradicionales tienen un problema de memoria corta debido al descenso del gradiente, Arana (2021) afirma que para mitigar el problema de short-term memory, aparece un nuevo tipo de celda de memoria que sí es capaz de extraer patrones de secuencias de mayor longitud. Según Gaset (2022), las celdas LSTM permiten a las RNN recordar sus entradas durante un largo

período de tiempo, debido a que los modelos LSTM contienen su información en la memoria, que puede considerarse similar a la memoria de una computadora, en el sentido que una neurona de una LSTM puede leer, escribir y borrar información de su memoria,

En este proyecto se realiza una revisión exhaustiva de la literatura existente sobre el uso de redes neuronales recurrentes (RNNs) para el pronóstico de la demanda. Se analizan los trabajos de autores como Almeyda (2022), Llerena (2022), Blanco et al. (2023), Ricce (2021), De La Cruz (2020) y Martínez y Ojeda (2018), quienes han aplicado con éxito las Redes Neuronales para abordar problemas de pronóstico de demanda.

En este sentido, el objetivo de este proyecto es diseñar un modelo dinámico para el pronóstico de la demanda en una empresa productora de alimentos. La empresa en estudio es multinacional, está ubicada en el estado Carabobo, Venezuela y se dedica a la producción de alimentos como salsa a base de tomate, salsas negras, compotas para bebés, salsas a base de mostaza, pastas, jugos, bebidas gaseosas y dulces en distintos formatos.

METODOLOGÍA

El presente trabajo es un proyecto factible, con un diseño de campo y metodología descriptiva, centrado en una unidad de análisis.

Como primera fase se realizó un diagnóstico del proceso actual de pronóstico en el departamento de planificación de la producción en la empresa de alimentos. El mismo inicia con el pronóstico de la demanda, el cual debe restringir el Plan Agregado enviado por el equipo de mercadeo, evaluando tanto los insumos como los factores que influyen en la producción. Al desagregar el mencionado Plan, se garantiza tener disponibilidad de la materia prima cuyo proceso de abastecimiento comienza cuando el equipo de planificación de la producción envía el pronóstico al equipo de compras, para evitar que el tiempo de respuesta del proveedor influya en la disponibilidad a tiempo del producto en planta. El proveedor debe enviar la materia prima a planta para cumplir el protocolo de recepción de MP; donde el equipo de Calidad toma muestras de los insumos y recursos para realizar pruebas fisicoquímicas y dimensionales, respectivamente, al ser aprobados son introducidos al almacén de MP.

El proceso de producción empieza al tener la demanda desagregada, la disponibilidad de MOD, disponibilidad de líneas y MP, el planificador de la producción envía las órdenes de trabajo al almacén de MP, donde al ser recibidas son identificadas las ubicaciones de los recursos e insumos para luego ser enviados al pesado y contabilizados de acuerdo con el FIFO y el FEFO. Seguidamente, son enviados al área

de predespacho de producción donde quedan a la espera del inicio del proceso producción.

En 2023, la empresa comenzó a implementar modelos matemáticos para el pronóstico de la demanda, la cual se centra en modelos estadísticos que consideran patrones lineales y en su mayoría horizontales, entre los cuales se encuentran la suavización exponencial simple y el promedio móvil ponderado.

Se utilizó el diagrama de Pareto para segmentar la problemática y determinar los productos que arrojaron mayor diferencia entre las ventas reales y lo pronosticado durante el año 2023, basado en el aporte de Ricce (2021) que utilizó el principio de Pareto para establecer los clientes que generan los mayores beneficios y concentrar el estudio en éstos (Ver tabla 1).

Tabla 1. *Diferencia entre demanda pronosticada y ventas reales de la empresa*

Causas	Diferencia	Acumulado
Demanda Sobreestimada	857.589	70 %
Demanda Subestimada	160.146	83 %
Falla de línea	64.898	88 %
Disponibilidad de materia prima	59.449	93 %
Prioridad otro SKU	56.984	98 %
Proyección de inventario inicial	24.627	100 %
Total	**1.223.693**	

En la Tabla 1, se observa el 80% de las causas es decir 1.017.735 cajas provienen de demanda sobreestimada y subestimada. distribuido en las categorías que se muestran en la tabla 2.

Tabla 2. *Categorías, Demanda Sobreestimada y Subestimada*

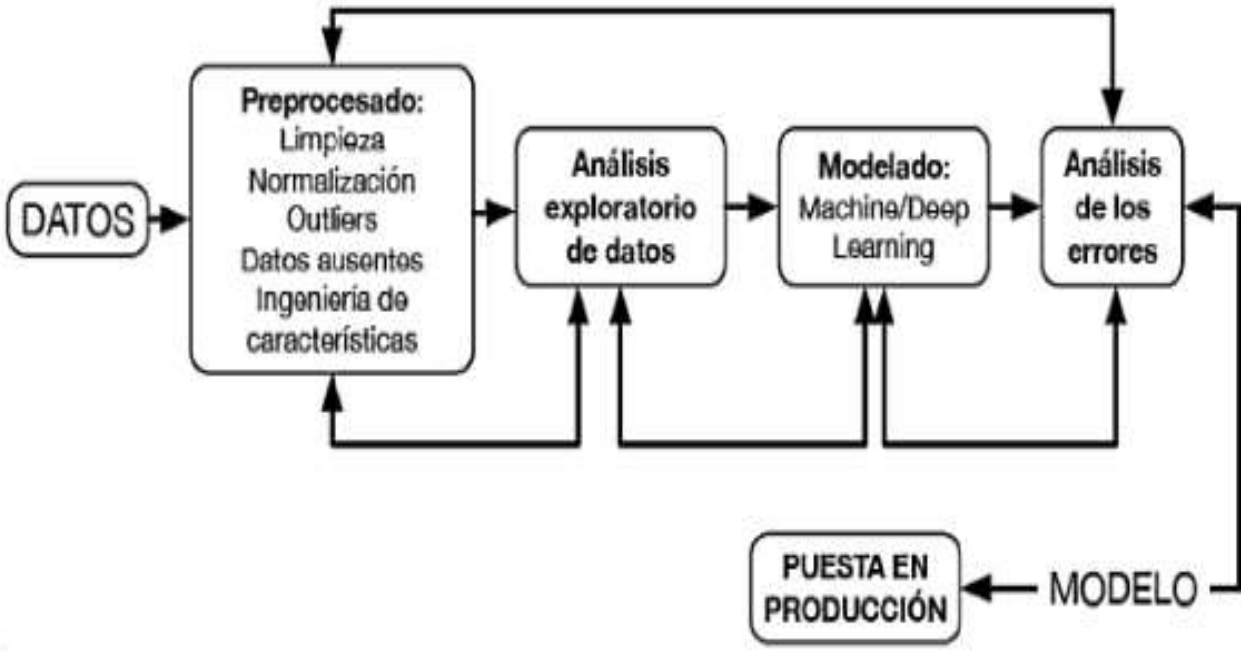
Categoría	Diferencia	Acumulado
Tomate	430.280	42 %
Bebés	236.793	69 %
Mostaza	65.566	77 %
Salsas para condimentar	52.594	83 %
Tomate condimentado	52.510	89 %
Pastas	40.559	94 %
Jugos	34.279	98 %
Bebidas gaseosas	20.595	100 %
Dulces	922	100 %
Total	**1.017.735**	

El 80% de las causas, el cual se representa por las categorías tomate, bebés, mostaza y salsas para condimentar, y con el ABC de ventas se seleccionaron los productos a estudiar, de los cuales resultaron 35 productos.

Para el análisis de datos, diseño y el desarrollo de la RNN se emplea el entorno Notebook Google Colab, lenguaje de programación Phyton, las bibliotecas Tensor Flow, Keras, Scikit-Learn, Numpy, Scipy, Statsmodels, Pandas,

En la fase dos donde se identificaron las variables que impactan en el pronóstico de la demanda se tomaron como base las etapas del análisis de datos como es muestra en la figura 1(Soria et al. 2023).

Figura 1. *Etapas del Análisis de Datos*



Para la normalización de los datos, la base fue tomada de Flores et. al (2021) para establecer las pruebas de hipótesis, quien recomienda utilizar la prueba de Shapiro-Wilk para muestras menores a 50 datos y para muestras mayores a 50 la prueba Kolmogórov-Smirnov en su corrección Lilliefors, las cuales tienen como hipótesis nula que los datos se comportan de forma normal y ésta se rechaza si el valor p es menor a alfa establecido como 0,05 es decir un 95% de confianza.

Para el tratamiento de los datos atípicos se toma como base el aporte de Angarita (2023), quien propone tres alternativas: mediante asimetría y curtosis, mediante rango intercuartil y mediante clustering, seleccionando el rango intercuartil, debido a que éste incluye implícitamente el análisis de asimetría y curtosis, además que no involucra tanto desarrollo algorítmico como el clustering.

En cuanto al manejo de datos atípicos, se considera el aporte de Gallo et al. (2021),

quienes proponen sustituir valores atípicos con un promedio de k valores vecinos, pronosticar los valores con un método predictivo y eliminar las observaciones. Se seleccionó la sustitución de valores atípicos con un promedio de k valores vecinos con el uso del algoritmo K-NN de Python que consigue el K óptimo para cada producto, debido a que pronosticar es uno de los objetivos de la investigación y se necesita la secuencia de ellos, por lo que eliminar datos no sería recomendable. La resultante para esta etapa fue el siguiente procedimiento:

1. El parámetro de interés: demanda del producto, Hipótesis nula (Ho): la demanda del producto sigue una distribución normal e Hipótesis alternativa (H1): la demanda del producto no sigue una distribución normal.
2. Nivel de significancia α=0,05 con una confianza del 95 %.
3. Estadístico de prueba
4. Se rechaza la hipótesis nula si el valor $p < \alpha$
5. Si los datos no son normales se procede a identificar los datos atípicos con el uso del rango intercuartil.
6. Los datos atípicos son sustituidos mediante el promedio de los k vecinos más cercanos.
7. Se vuelve a probar la hipótesis de normalidad.

Para comprender el comportamiento de la demanda de cada producto a lo largo del tiempo, se realizó un análisis de sus series temporales. De acuerdo con Villavicencio (2010), una serie temporal es una secuencia de observaciones, medidos en determinados momentos del tiempo, ordenados cronológicamente y, espaciados entre sí de manera uniforme, así los datos usualmente son dependientes entre sí. El mismo autor clasifica las series de tiempo en estacionaria y no estacionaria, se considera estacionaria si sus propiedades estadísticas, como la media y la varianza, permanecen constantes a lo largo del tiempo.

Para determinar si una serie de tiempo es estacionaria, se utiliza la prueba de Dickey-Fuller explicada por Asteriou (2002). Esta prueba estadística evalúa la presencia de una raíz unitaria en la serie. La hipótesis nula de la prueba es que la serie tiene una raíz unitaria, lo que significa que no es estacionaria. Si el valor p de la prueba es menor que el nivel de significancia (α), establecido en 0,05, se rechaza la hipótesis nula y se concluye que la serie es estacionaria.

En el caso de series de tiempo no estacionarias, se debe aplicar un proceso de desestacionalización antes de realizar la descomposición. Para ello, se utiliza el método multiplicativo, propuesto por Chase y Jacobs (2009). Los modelos de descomposición son: aditivo y multiplicativo, estos dividen las series de tiempo en sus componentes: tendencia, estacionalidad y residuos. Se siguió el siguiente procedimiento:

1. Se prueba la estacionalidad mediante Dickey-fuller, Hipótesis nula (H_0): la demanda del producto no estacionaria e Hipótesis alternativa

(H_1): la demanda del producto es estacionaria

2. Nivel de significancia α=0,05 con una confianza del 95%.

3. Estadístico de prueba es análogo al de una t student pero sigue una distribución distinta por lo que se utiliza el valor crítico obtenido de la tabla Dickey Fuller

4. Se rechaza la Ho si Estadístico < valor crítico

5. Si la serie de tiempo tiene más de 24 datos es decir más de 2 años se realiza la descomposición de la serie de tiempo.

6. Si la serie de tiempo es estacionaria se usa el método aditivo sino multiplicativo.

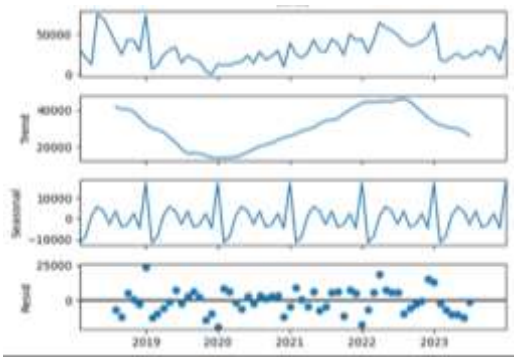
7. Mediante el método visual se observa la tendencia y la estacionalidad.

El análisis de las series de tiempo permitió identificar diferentes tipos de estacionalidades y tendencias en los datos. Se encontraron estacionalidades anuales, semestrales, cuatrimestrales y trimestrales, lo que indica que la demanda de algunos productos presenta patrones recurrentes en diferentes intervalos de tiempo. También se observaron tendencias tanto lineales como no lineales en los datos (Ver figuras 2, 3, 4, 5 y 6).

Tomate 400 g

Comportamiento: Normal, estacionario, tendencia no lineal y estacionalidad semestral.

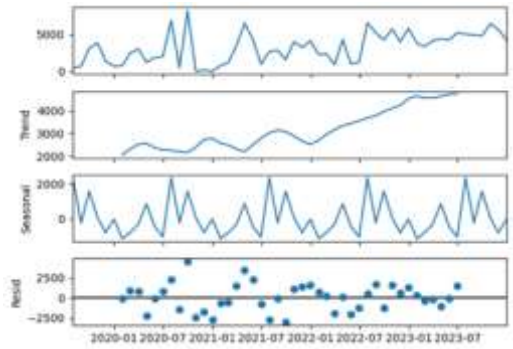
Figura 2. *Descomposición Tomate 400 g*



Tomate clase B 400 g

Comportamiento: Normal, estacionario, tendencia no lineal y estacionalidad semestral.

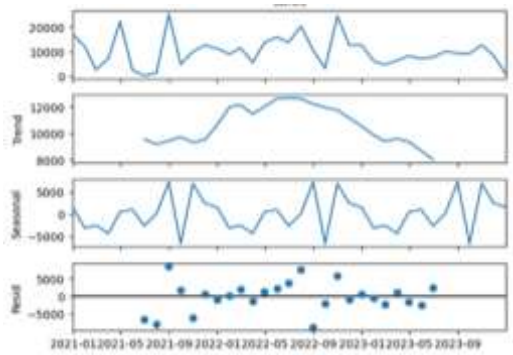
Figura 3. *Descomposición Tomate clase B 400g*



Bebés Pera 200 g

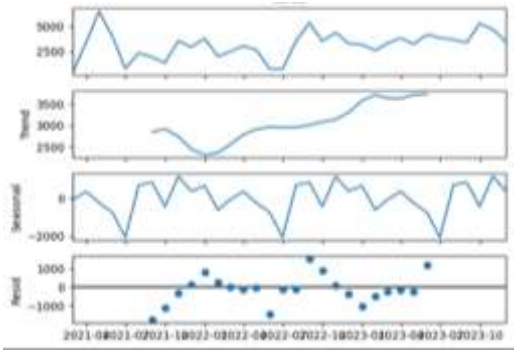
Comportamiento: Normal, estacionario, tendencia no lineal y estacionalidad cuatrimestral.

Figura 4. *Descomposición Bebés Pera 200 g*



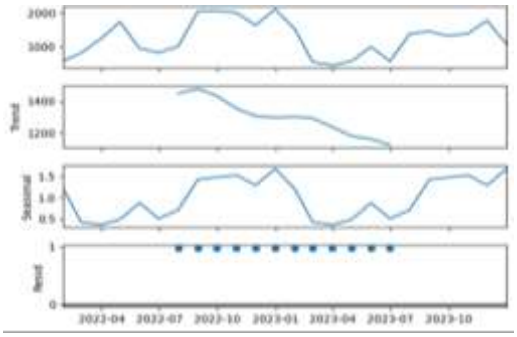
Bebés Plástico Manzana 100 g
Comportamiento: Normal, estacionario, aproximado lineal y trimestral.

Figura 5. *Descomposición Bebés Plástico Manzana 100 g*



Bebés Plástico Uva 200 g
Comportamiento: Normal, no estacionario, aproximado lineal y trimestral.

Figura 6. *Descomposición Bebés Plástico Uva 200 g*



Se analizaron los datos de 35 productos para determinar la normalidad y estacionariedad. De estos productos, solo tres no presentaron datos atípicos y no se pudieron normalizar: Tomate clase B 200 g, Bebés piña 200 g y uva 100 g. Sin embargo, en nueve productos sí se identificaron valores atípicos, pero no fue posible normalizar los datos: Tomate 200 g, Bebés manzana 200 g, Manzana 100 g, fresa 200 g, fresa 100 g, piña 100 g, durazno 100g,uva 200g y durazno 200 g.

Adicionalmente, se encontró que siete de los 35 productos no son estacionarios: Mostaza 3000 g, 100 g y 10 g, Bebés plástico Piña 100 g, durazno 100 g, uva 200 g y Salsa 100 g, los otros 28 productos presentaban series estacionarias, lo que significa que sus propiedades estadísticas (media, varianza) permanecían constantes en el tiempo.

Continuando el proceso investigativo, en la fase tres se elaboró la propuesta, en la cual se diseñó una red neuronal multicapa de celdas LSTM cuyos hiper parámetros del modelo fueron establecidos tomando como base el aporte de Blanco et al. (2023), el cual propone probar con 50 neuronas para cada capa de celdas LTSM, en su trabajo de grado la problemática era más compleja teniendo una arquitectura de 5 capas LSTM con una capa de dropout (dilución) del 20%. En la tabla 3 se presentan las características de la RNN diseñada.

En este proyecto, se consideró como punto de partida para el entrenamiento, establecer los hiper parámetros de éste usando como referencia el producto con mayor cantidad de datos ya que es considerado el que contiene mayor información. Se crearon nuevos conjuntos de datos que se ajusten al tamaño del producto de mayor cantidad de datos utilizando la herramienta bootstrap, lo que permite obtener una estimación más robusta del rendimiento del modelo y realizar ajustes más precisos a los otros productos.

Tabla 3. *Hiper Parámetros del Modelo*

	Características
Vector de entrada	Es de dos dimensiones, donde la primera dimensión representa el número de pasos de tiempo (en este caso, 1) y la segunda dimensión representa la característica (valor de venta).
Capa LSTM 1	Procesa la secuencia de datos escalados. Contiene 50 unidades ocultas. Devuelve la salida para cada paso de tiempo.
Dropout	Con una tasa de deserción del 20% para prevenir el sobreajuste.
Capa LSTM 2	Procesa la salida de la capa LSTM 1. Contiene 50 unidades ocultas. No devuelve la salida para cada paso de tiempo.
Capa Densa	La capa densa final transforma la representación vectorial compleja de la capa LSTM 2 en un valor escalar. Contiene una sola neurona. Genera la predicción de ventas para el siguiente paso de tiempo.
Función de activación	tanh

Cabe destacar, que de acuerdo con Núñez (2000), la idea principal de los métodos bootstrap es el remuestreo a partir de los datos originales ya sea en forma directa o vía un modelo ajustado; con la finalidad de obtener muestras replicadas a partir de las cuales se pueda evaluar la variabilidad de las cantidades de interés sin incurrir en errores de cálculos analíticos.

El entrenamiento se centró en establecer un número de épocas que arroje un error del 5% para evitar el sobreajuste, Nanclares (2021) explica que las épocas son el número de iteraciones al conjunto de entrenamiento, se probó con 5, 20, 40 y 80 épocas, los demás hiper parámetros de entrenamiento se tomaron del aporte de Blanco et al. (2023).

En la figura 7 se puede observar que el entrenamiento con 5 épocas está incompleto. En la figura 8 se puede observar que el entrenamiento con 80 épocas da un resultado de pronóstico mucho más preciso, obteniendo en la figura 9, la curva de validación para las 80 épocas.

Figura 7. 5 *Épocas con Producto de Mayor Cantidad de Datos*

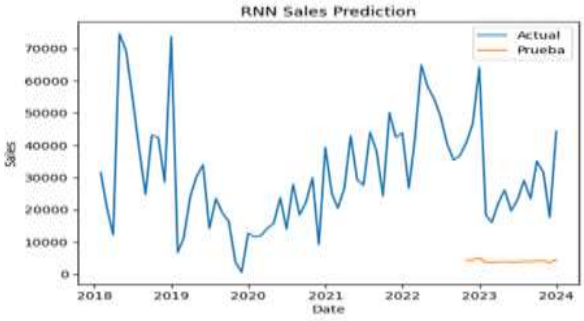


Figura 8. *80 Épocas con Producto de Mayor Cantidad de Datos*

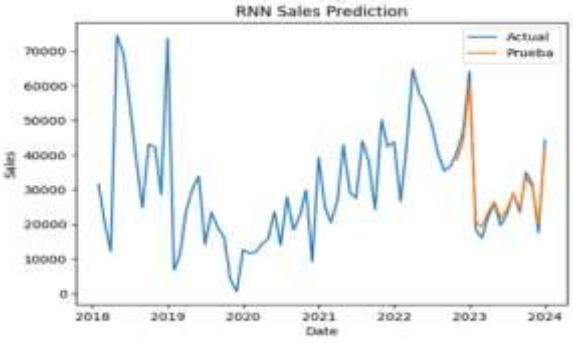
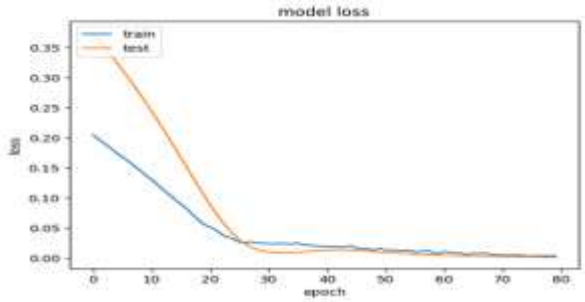


Figura 9. *Curva de validación para 80 épocas*



Se observa que la curva de validación comienza a aumentar después de la época 80. Esto indica que el modelo está comenzando a sobre ajustarse a los datos de entrenamiento, entonces se detienen las pruebas. La línea roja representa el MSE del modelo durante el entrenamiento, se observa que la curva desciende gradualmente a lo largo de las épocas, lo que indica que el modelo mejora su precisión con el entrenamiento, la línea azul representa el MSE del modelo en un conjunto de datos de validación independiente. Este conjunto de datos no se utiliza para entrenar el modelo, sino para evaluar su generalización a datos nuevos. Se observa que la curva de validación sigue una tendencia similar a la curva de entrenamiento y el punto de intersección está en un punto bajo del MSE, lo que indica que el modelo ha aprendido patrones generales sin sobre ajustarse a los datos específicos del entrenamiento.

En general, se estableció el siguiente procedimiento para crear la red neuronal multicapa de celdas LSTM :

1. Escalar los valores de entrada a un rango entre 0 y 1, para que las neuronas aprendan de manera más uniforme y eficiente, acelerando la convergencia del entrenamiento y mejorando la precisión del modelo.
2. Identificar vector de entrada.
3. Seleccionar el 80 % de los datos escalados como conjunto de entrenamiento, este conjunto se utiliza para ajustar los pesos de la red neuronal durante el proceso de entrenamiento.
4. El 20 % restante de los datos escalados se reserva como conjunto de prueba, este conjunto se utilizará para evaluar el rendimiento del modelo una vez entrenado.
5. Establecer que sea secuencial.
6. Establecer las cantidades de capas LSTM que contiene el modelo y para cada una se fija la cantidad de neuronas.
7. Asignar una capa Dropout de 20 % de deserción, para ayudar a prevenir el sobreajuste del modelo al eliminar aleatoriamente un porcentaje de las neuronas durante el entrenamiento.
8. Agregar una capa densa final con una neurona, esta capa transforma la representación vectorial compleja en un valor escalar, que representa la predicción de ventas para el siguiente paso de tiempo.
9. Entrenar el modelo.

La última fase consistió en la validación de la propuesta mediante un análisis estadístico, ejecutando las corridas del modelo propuesto con los pronósticos sugeridos y los datos de la situación actual presentando gráficas donde se comparen los resultados obtenidos con el modelo propuesto y el modelo actual.

Soria (2023) explica que para validar un modelo de machine learning o deep

learning hay que analizar la calidad de su rendimiento mediante diferentes medidas de error que se pueden usar para comparar con otros modelos y elegir finalmente el que mejor se adecue al problema de los datos, en este caso es importante mencionar el sesgo y la varianza.

Para establecer la confiabilidad de los resultados del modelo se toma como base el aporte de Sipper (1998), el cual demostró que DAM/0,8 es una buena estimación de la desviación estándar aun para una distribución no normal. Su estudio empírico mostró que para distribuciones uniforme, exponencial o triangular, la constante sólo varía entre 0,74 y 0,86, por lo que 0,8 es igualmente una buena aproximación (Ver tabla 4).

Tabla 4. *Interpretación de los resultados*

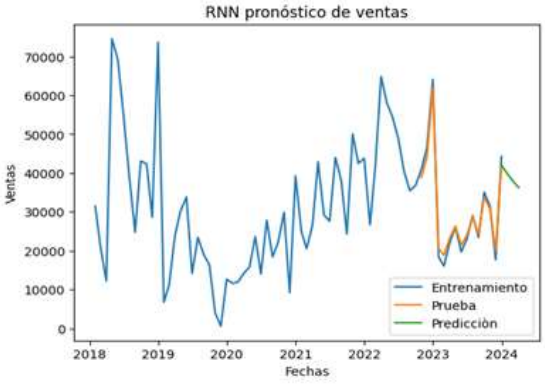
Parámetro	Interpretación
Error medio (EM)	EM > 0 en promedio, se subestimó la demanda
	EM = 0, hay error, pero el pronóstico este alrededor de la demanda
	EM < 0 En promedio se subestimó la demanda
Error cuadrático medio (ECM)	ECM > 0 hubo error
	ECM = 0, la demanda es igual al pronóstico
	ECM < 0 imposible
Desviación media absoluta (DMA)	DMA > 0 hubo error
	DMA = 0, la demanda es igual al pronóstico
	DMA < 0 imposible
Precisión	Un valor de RMSE bajo indica que el modelo de pronóstico es preciso y genera predicciones cercanas a los valores reales.

RESULTADOS

Se presentan los resultados obtenidos del modelo de red neuronal multicapa recurrente con celdas Long Short-Term Memory (LSTM) desarrollado para el pronóstico de la demanda. Se analizará el rendimiento del modelo en términos de precisión y se compararon los resultados del modelo LSTM con los pronósticos obtenidos mediante el método actual, destacando las mejoras y avances logrados (Ver figuras 10, 11, 12, 13, 14, 15 y 16).

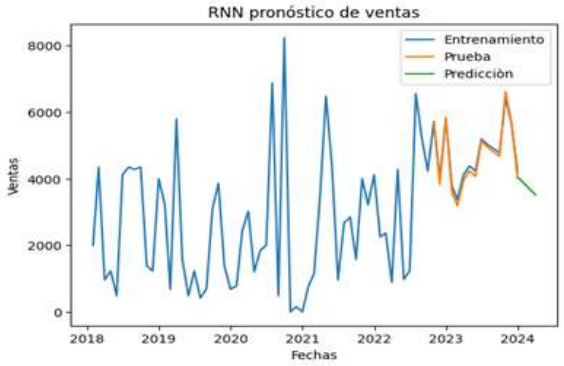
Tomate 400 g: Precisión: 94,84 %.

Figura 10. *Pronóstico Tomate 400 g*



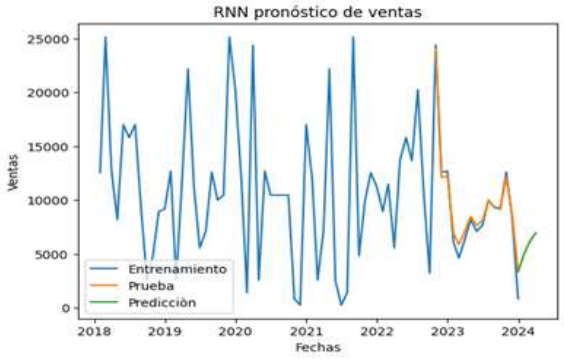
Tomate clase B 400 g: Precisión: 97,53%.

Figura 11. *Pronóstico Tomate clase B 400 g*



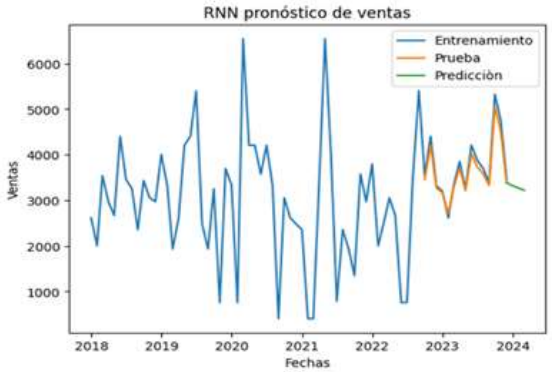
Bebés Pera 200 g: Precisión 93,63 %.

Figura 12. *Pronóstico Bebés Pera 200 g*



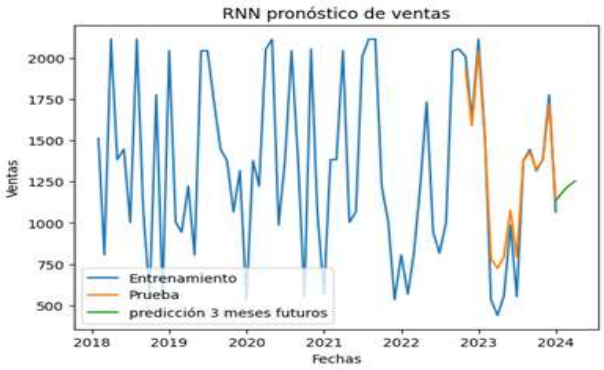
Bebés Plástico Manzana 100 g: Precisión 96,76 %.

Figura 13. *Pronóstico Bebés Plástico Manzana 100 g*



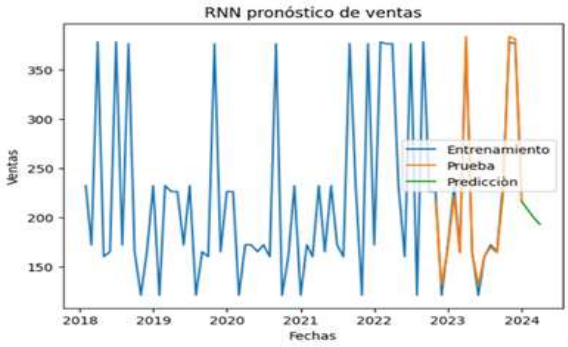
Bebés Plástico Uva 200 g: Precisión: 91,32%.

Figura 14. Pronóstico Bebés Plástico Uva 200g



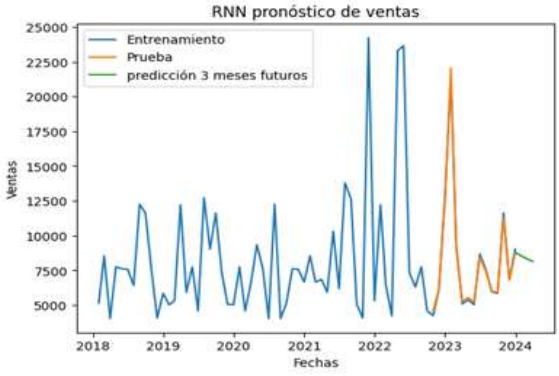
Mostaza 3000 g: Precisión: 97,47 %, normal, menor a 24 meses, no estacionario.

Figura 15. *Mostaza 3000 g*



Tomate 200 g: Precisión: 94,49 %, no normal.

Figura 16. Pronóstico Tomate 200 g



Los resultados de pronóstico de la LSTM-RNN de valores futuros a 3 meses (De Enero a Marzo de 2024) tomando como último dato de todas las series de tiempo el valor de las ventas de Diciembre de 2023.

Es importante considerar el intervalo de confianza al 95%, el cual indica que existe una cierta variabilidad en las predicciones. Esto significa que la demanda real podría ser mayor o menor que los valores pronosticados dentro de este rango considerando el aporte de Sipper (1998). Seguidamente los resultados de los productos mencionados anteriormente (Ver tabla 5).

Por último,en las figuras 17, 18, 19, 20, 21, 22 y 23 se muestran los resultados de la comparación estadística del modelo propuesto y el modelo actual utilizado por la empresa para determinar cuál método proporciona mejor resultado usando para esto la desviación media absoluta (DMA).

Tabla 5. *Pronóstico 3 meses futuros*

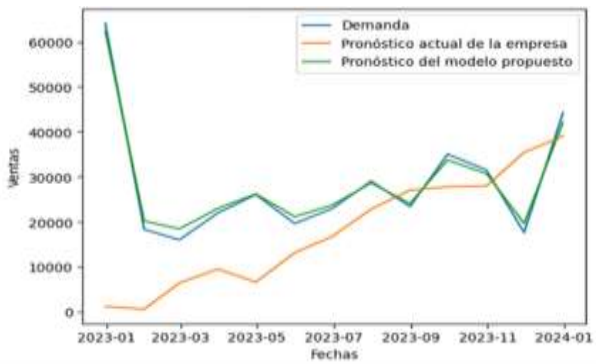
Nombre	Enero	Febrero	Marzo	Normal	DMA	Intervalo de Confianza 95 %
Tomate 400 g	39.841,60	37.937,33	36.248,29	X	1.574,90	3.858,60
Tomate 200 g	8.551,85	8.336,96	8.138,10	X	223,70	548,10
Tomate clase B 400g	3.871,89	3.696,06	3.514,87	X	117,80	288,50
Bebés Pera 200 g	4.993,10	6.121,69	6.932,05	X	593,30	1.453,50
Bebés plástico Manzana 100 g	3.319,88	3.266,28	3.222,68	X	121,40	297,50
Bebés plástico Uva 200 g	1.184,83	1.123,26	1.252,77		97,60	
Mostaza 3000 g	207,61	202,60	192,83	X	5,50	13,60

Tomate 400 g

DMA modelo actual: 13.716,7814 cajas

DMA modelo propuesto:1.285,3113 cajas.

Figura 17. *Modelo Actual vs Modelo Propuesto Tomate 400 g*

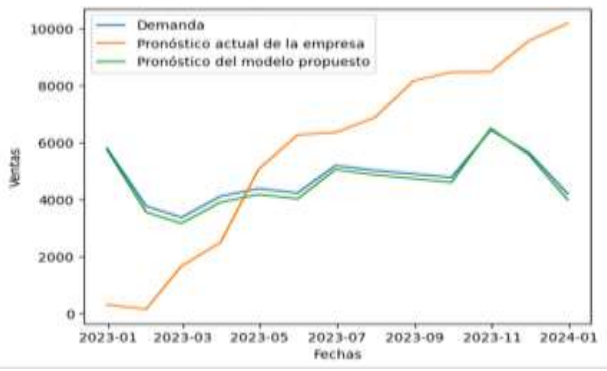


Tomate clase B 400 g

DMA modelo actual: 2.848,0271 cajas

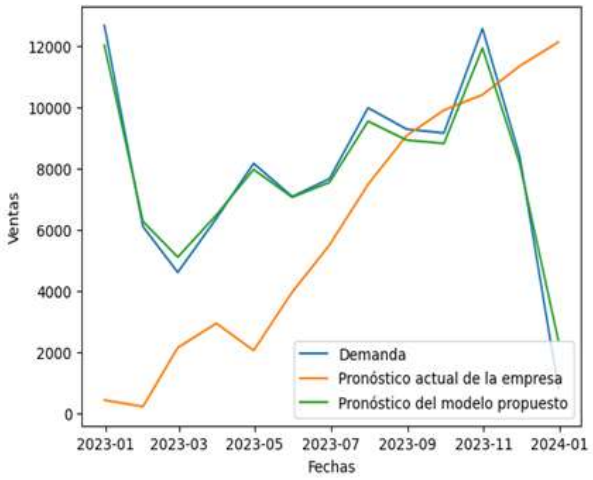
DMA modelo propuesto: 159,7510 cajas

Figura 18. *Modelo Actual vs Modelo Propuesto Tomate clase B*



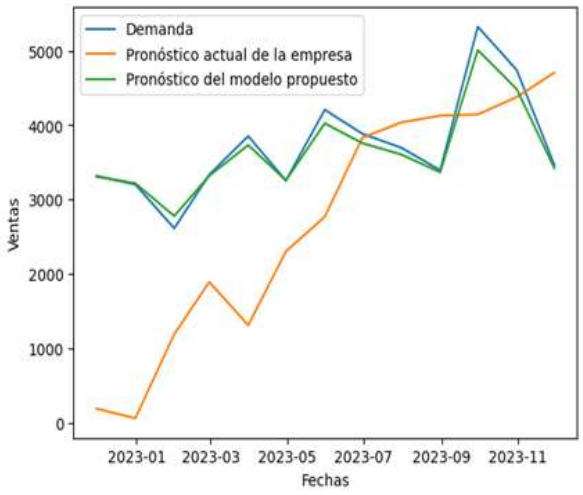
Bebés Pera 200 g
DMA modelo actual: 4.260,2321 cajas
DMA modelo propuesto: 411,2157.cajas

Figura 19. *Modelo Actual vs Modelo Propuesto Bebés Pera 200 g*



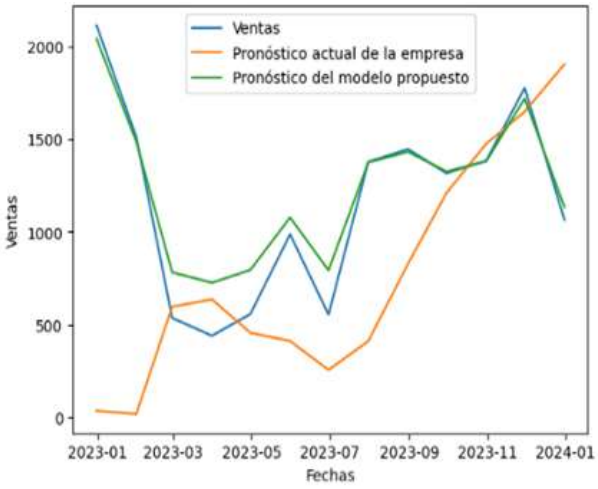
Bebés Plástico Manzana 100 g
DMA modelo actual: 1.382,7862 cajas
DMA modelo propuesto: 105,5871 cajas

Figura 20. *Modelo Actual vs Modelo Propuesto Bebés plástico manzana 100 g*



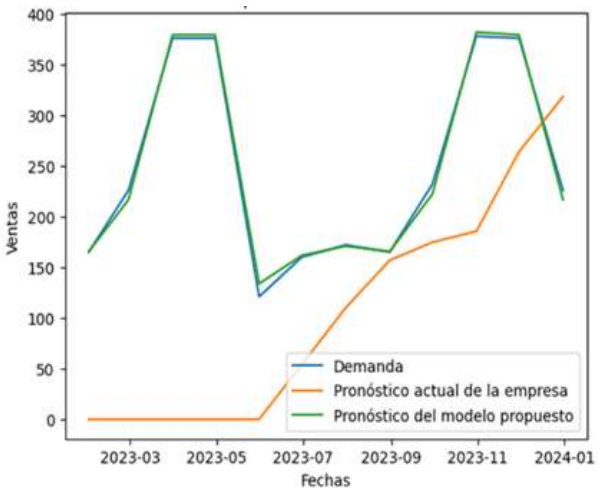
Bebés Plástico Uva 200 g
DMA modelo actual: 580,9 und
DMA modelo propuesto: 104.und

Figura 21. *Modelo Actual vs Modelo Propuesto Bebés plástico uva 200 g*



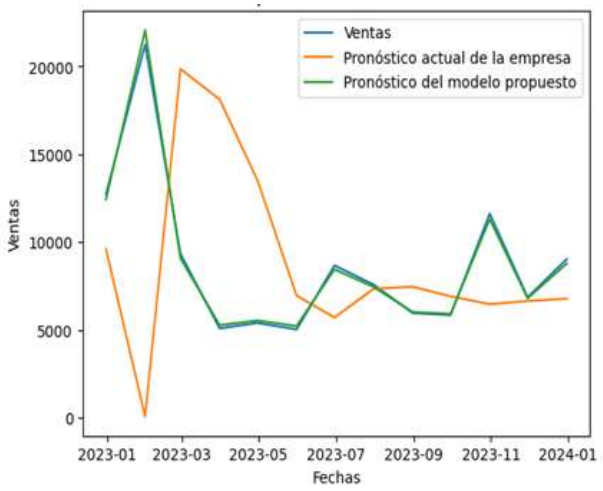
Mostaza 3000 g
DMA modelo actual: 157,86 cajas
DMA modelo propuesto: 4,91 cajas

Figura 22. *Modelo Actual vs Modelo Propuesto Mostaza 3000 g*



Tomate 200 g
DMA modelo actual: 5.481,8207 cajas
DMA modelo propuesto: 297, 65 cajas

Figura 22. *Modelo Actual vs Modelo Propuesto Tomate 200 g*



Discusión de los resultados

En el pronóstico de 3 meses futuros, se consiguieron 5 de los 35 productos estudiados cuyo comportamiento de los errores no es normal, pero basándose en la contribución de Sipper. (1998) y luego de evaluar con prueba de hipótesis de comportamiento uniforme aceptando ésta para los 5 productos se confirma que los resultados son confiables. Los mismos fueron Tomate clase B 200 g, Bebés piña 200 g, piña 100, uva 200 g y Bebés plástico uva 200 g .

Se utiliza la desviación media absoluta DMA para comparar cuál método proporciona mejores resultados como propone Soria (2023) resultado para todos los productos que el método propuesto es mejor ya que para todos se consiguió una DMA menor que el método actual utilizado por la empresa.

Se observa que para todos los productos disminuyó la subestimación y la sobre estimación de la demanda de acuerdo con la desviación media absoluta, obteniendo una precisión promedio del modelo propuesto de 94,11 %, en comparación del método actual utilizado por la empresa que tiene en promedio 30,22 % de precisión, mejorando en promedio, un 63,89 %.

CONCLUSIONES

En conclusión, en el análisis exhaustivo, se evaluó el proceso actual de pronóstico de demanda en la empresa en estudio utilizando el principio de Pareto, cuya finalidad era identificar las áreas que generan el mayor impacto en la precisión del pronóstico y enfocar los esfuerzos de mejora en esas áreas críticas. El análisis de Pareto reveló que un pequeño porcentaje de los productos (alrededor del 20%) genera una gran parte del error total de pronóstico (alrededor del 80%) con un total de la diferencia entre el pronóstico y las ventas reales en cajas de 1.223.693, de las cuales 1.017.735 pertenecen a la categoría tomate, bebés, mostaza y condimentados, resultando como principal problemática la demanda sobreestimada y subestimada.

Así mismo, se lograron identificar las variables mediante el uso de herramientas estadísticas tradicionales para conocer el comportamiento de la demanda de los

distintos productos, dándonos certeza que algunos de los datos tienen una distribución conocida como la normal, confirmando ésto con pruebas de hipótesis como lo son la prueba de Lilliefors y la de Saphiro Wilk con un 95% de confianza; así como identificar los datos atípicos, los cuales se alejan del comportamiento normal de los productos. Para ello, se utilizaron herramientas como el método de caja o el rango intercuartil y posteriormente sustituir estos datos para no perder la secuencia de las datas históricas, utilizando el método de k vecinos más cercano, logrando normalizar algunos de los comportamientos de las demandas de los productos, aunque también otros no se lograron normalizar. Con el uso de la prueba de hipótesis de Dickey-Fuller se identificaron aquellas series de tiempo que se comportan de forma estacionaria y no estacionaria. Por ser importante para el estudio, también se utilizó la descomposición clásica con el análisis visual, de lo cual se identificaron las tendencias tanto lineales como no lineales y las estacionalidades: trimensuales, cuatrimestrales, semestrales y anuales, esto con la finalidad, de determinar el comportamiento de los datos para ser introducidos al modelo.

Las redes neuronales recurrentes (RNN), particularmente las de tipo Long Short-Term Memory (LSTM), han demostrado ser herramientas poderosas para modelar secuencias de datos y capturar dependencias a largo plazo, estas se combinaron con métodos estadísticos tradicionales como el bootstrap hicieron una herramienta que se puede utilizar para todos los productos de la empresa de alimentos, con una precisión mayor al 80% en todos los casos. Además, se concluye que el método propuesto tiene mejor ajuste que el método actual en términos de errores de pronóstico como el error medio que nos indicó, demandas sobreestimada y subestimada, pero en menor cantidad que el método actual de la empresa, el error cuadrático y error absoluto que para todos los productos indicó que el método propuesto es más preciso que el método actual.

Se concluye que la implementación de un modelo multicapa de redes recurrentes de celdas LSTM puede ser una solución efectiva para mejorar la precisión del pronóstico de la demanda y mantenerla en el tiempo. Este tipo de modelo permite capturar mejor las relaciones temporales y patrones complejos en los datos de demanda, lo que ayuda a predecir con mayor precisión las necesidades futuras de producción; esta predicción resulta crucial para la toma de decisiones en el departamento de planificación de la producción, debido a que el modelo ajusta los pesos para obtener una precisión constante en el tiempo y confiable para optimizar los recursos y evitar problemas de sobre o sub producción lo que hace que se garantice la eficiencia operativa.

REFERENCIAS

Almeyda, E. (2022). *Pronóstico de la demanda internacional del banano orgánico de Perú usando algoritmos de Machine Learning*. [Tesis Doctoral]. Universidad de Piura, Perú. https://hdl.handle.net/11042/5718

Angarita, A. (2023). *¿Cómo detectar y tratar los outliers?*. Siepsi. https://siepsi.com.co/2023/06/16/como-detectar-y-tratar-los-outliers/

Arana, C. (2021). *Redes neuronales recurrentes: Análisis de los modelos especializados en datos secuenciales, Serie Documentos de Trabajo, No. 797*. Universidad del Centro de Estudios Macroeconómicos de Argentina (UCEMA). https://hdl.handle.net/10419/238422

Asteriou, D. (2002). *Notas sobre Análisis de Series de Tiempo: Estacionariedad, Integración y Cointegración*. Webdelprofesor, Universidad de Los Andes, ULA. https://webdelprofesor.ula.ve/economia/hmata/Notas/Notas%20sobre%20Analisis%20de%20Series%20de%20Tiempo.pdf

Blanco, I., García, S. J., & Remesal, Á. (2023). Aprendizaje profundo para series temporales en finanzas: aplicación al factor momentum. En Funcas (Ed.), *Análisis financiero y big data* (Capítulo V, pp. 123–148). Funcas. https://www.funcas.es/wp-content/uploads/2023/05/Analisis-financiero-y-big-data_Capitulo-V.pdf

Chase, R. y Jacobs, R. (2009). *Administración de operaciones Producción y cadena de suministros*. McGraw-Hill / Interamericana.

Del Castillo, S. (2023). *Logística 4.0: Innovación y eficiencia en la cadena de suministro, 1era Edición*. Doxa Edition.

De La Cruz, C. (2020). *Previsión De La Demanda Mediante Deep Learning* [Trabajo fin de Máster, Universidad de Sevilla]. idUS. Depósito de Investigación de la Universidad de Sevilla. https://hdl.handle.net/11441/106992

Flores, C. y Flores, K. (2021). Pruebas para comprobar la normalidad de los datos en procesos productivos: Anderson- Darling, Ryan- Joiner, Shapiro- Wilk y Kolmogorov-Smirnov. *Societas. Revista de Ciencias Sociales y Humanísticas*, 23(2), 83-106. https://portal.amelica.org/ameli/journal/341/3412237018/3412237018.pdf

Gallo, A.; Pérez, F.; Salinas, D. (2021). Minería de Datos y Proyección a Corto Plazo de la Demanda de Potencia en el Sistema Eléctrico Ecuatoriano. *Revista Técnica "energía"*, 18(1), 72-85. https://revistaenergia.cenace.gob.ec/index.php/cenace/article/view/461/644

Gaset, C. (2022). *Redes neuronales recurrentes aplicadas a series de tiempo: predicción de pasajeros de rutas regionales en Argentina*. [Tesis Master, Universidad Torcuato di Tella]. Repositorio Digital Universidad Torcuato Di Tella. https://repositorio.utdt.edu/handle/20.500.13098/11575

Llerena, R. (2022). *Propuesta de un Modelo Matemático aplicado al pronóstico de Producción utilizando Redes Neuronales Artificiales empleado en una Fábrica de Cajas de Seguridad Modelo 39SS.* [Tesis Maestría, Universidad Estadal de Milagro]. Repositorio Institucional de la Universidad Estatal de Milagro UNEMI. http://repositorio.unemi.edu.ec//handle/123456789/6423

Martínez, D. y Ojeda, A. (2018). *Análisis de diferentes métodos de pronóstico en la cadena de suministro a través de un modelo dinámico de simulación.* [Trabajo Especial de Grado no publicado Universidad de Carabobo]. Universidad de Carabobo.

Nacelle, A. (2009). *Las redes neuronales artificiales: de la biología a los algoritmos de clasificación.* [Monografía]. Universidad de la República Uruguay.

Nanclares, J. (2021). *Redes Neuronales recurrentes para la predicción de series temporales.* [Trabajo de Grado, Universidad Autónoma de Madrid]. Repositorio Institucional UAM. https://repositorio.uam.es/server/api/core/bitstreams/8eeee004-ea2b-4b05-98f3-cdf40db99fb2/content

Núñez, G. (2000). Introducción al Bootstrap y sus aplicaciones. En *INEGI, Memorias del XIV Foro Nacional de Estadística* (pp. 81–86). Instituto Nacional de Estadística, Geografía e Informática.

Render et al., (2007). *Administración de la producción, Primera Edición*. Pearson Educación.

Ricce, E. (2021). *Pronóstico de demanda altamente variable e intermitente usando un modelo básico de Red Neuronal Artificial para disminuir el riesgo de rotura de stock de una compañía que abastece productos en Sudamérica.* [Trabajo de Grado, Universidad Peruana de Ciencias Aplicadas]. Repositorio Académico UPC. https://repositorioacademico.upc.edu.pe/handle/10757/657936

Rouhiainen, L. (2018). *Inteligencia Artificial 101 Cosas que debes saber hoy sobre nuestro futuro*. Planeta.

Sandoval, L. (2018). Algoritmos de aprendizaje automático para análisis y predicción de datos. *Revista Tecnológica*, 11, 36-40. http://www.redicces.org.sv/jspui/bitstream/10972/3626/1/Art6_RT2018.pdf

Sipper, D. (1998). *Planeación y control de la producción*. McGraw Hill Interamericana Editores S.A.

Soria, E., Sánchez-Montañés, M., Gamero, R., Castillo, B. y Cano, P. (2023). *Sistemas de aprendizaje automático*. Ra-Ma.

Villavicencio, J. (2010). Introducción a Series de Tiempo. Recuperado de http://www.estadisticas.gobierno.pr/iepr/LinkClick.aspx?fileticket=4_BxecUaZmg%3D

Autores

Normeviz Depablo Vizcaya,. Ingeniero Industrial, Universidad de Carabobo; Magister en Ingeniería Industrial, Universidad de Carabobo, Venezuela.

ORCID: https://orcid.org/0009-0008-7247-5693
Email: normevizdepablo@gmail.com

Paola Martínez Valero. Ingeniero Industrial, Universidad de Carabobo, Venezuela.

ORCID: https://orcid.org/0009-0006-2734-9036
Email: pao29mariana@gmail.com

Francisco José. Figueredo Lugo. Ingeniero Ingeniero Industrial. Profesor-Investigador, Departamento de Investigación de operaciones, Escuela de Ingeniería Industrial, Facultad de Ingeniería, Universidad de Carabobo, Venezuela.

ORCID: https://orcid.org/0009-0002-7857-4971

Email: ffigueredo@uc.edu.ve

Recibido: 30-03-2025 **Aceptado:** 29-06-2025