



UNIVERSIDAD DE CHILE

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

**EVALUACIÓN DE LA INCORPORACIÓN DE NUEVAS VARIABLES EN LA  
PREDICCIÓN DE DEMANDA PARA UNA EMPRESA DE ALIMENTOS**

**MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL**

**JORGE EDUARDO NARANJO SAN MARTÍN**

**PROFESOR GUÍA:  
NICOLÁS CISTERNAS GONZÁLEZ**

**MIEMBROS DE LA COMISIÓN:  
LORETO MARTÍNEZ GIMÉNEZ  
GASTÓN SUARÉZ CROTHERS**

**SANTIAGO DE CHILE  
2023**

**RESUMEN DE LA MEMORIA PARA OPTAR  
AL TÍTULO DE:** Ingeniero Civil Industrial  
**POR:** Jorge Eduardo Naranjo San Martín  
**FECHA:** 2023  
**PROFESOR GUÍA:** Nicolás Cisternas González

## **EVALUACIÓN DE LA INCORPORACIÓN DE NUEVAS VARIABLES EN LA PREDICCIÓN DE DEMANDA PARA UNA EMPRESA DE ALIMENTOS**

Agrosuper es una empresa chilena del sector alimentario de productos cárnicos, la cual alcanzó en 2022 un volumen de ventas de \$4.179 millones de dólares.

Dentro de la estructura de Agrosuper, se encuentra el área de planificación de demanda, donde la principal responsabilidad de esta área es la anticipación, estimación y gestión precisa de la demanda futura de los productos que la empresa ofrece. En su búsqueda constante por mejorar la precisión en sus pronósticos de demanda, esta área ha identificado una oportunidad de mejora que reside en la incorporación de nuevas variables en sus modelos de predicción, ya que en la actualidad solo están utilizando la demanda histórica para sus pronósticos de demanda.

Lograr una mayor precisión en los pronósticos de demanda beneficiaría al área de cadena de suministro en varios aspectos, ya que facilitaría una planificación de demanda más exacta, proporcionaría una gestión de inventario más eficaz y perfeccionaría la eficiencia operacional [7]. De manera conjunta, estas mejoras podrían traducirse en ahorros sustanciales de tiempo y recursos económicos.

Este proyecto se enfoca en esta oportunidad de mejora, con el objetivo de evaluar la precisión al introducir nuevas variables en la predicción de demanda, específicamente enfocada en el sector de productos procesados (elaborados y cecinas) de la empresa. Para ello, se propone desarrollar una red neuronal y variaciones de esta misma red utilizando las distintas combinaciones de variables. Para alcanzar este objetivo, se utilizó la metodología CRISP-DM, la cual se basa en un ciclo iterativo que comprende cinco fases: comprensión del negocio, comprensión de los datos, preparación de los datos, modelado y evaluación. Bajo esta metodología, se desarrolló una red neuronal LSTM (Long Short-Term Memory), que tiene la capacidad de aprender y recordar información a largo plazo [14]. También se consideraron variables internas de la empresa, las cuales fueron el precio, la facturación en kilos y la elasticidad de la demanda.

En el desarrollo de este proyecto, se estableció un caso base que consistente en una red LSTM que solo utiliza la demanda histórica. Adicionalmente, se generaron variantes de esta red que integran diversas combinaciones de las variables seleccionadas, con el fin de evaluar la precisión de los pronósticos comparados a este caso base. Tras evaluar la precisión de los pronósticos entregados por las diversas redes, se observa que, en general, indican una mayor precisión en comparación con el caso base. En particular, para el sector de productos elaborados, al considerar la variable del precio, se obtiene un pronóstico con una precisión promedio del 90,4% contra un caso base de 88,5%. Y para el sector de cecinas, utilizando las variables precio y elasticidad se tiene una precisión del 86,2% contra un caso base de 80,5%. De esto se deduce que, al incorporar nuevas variables, se logra una mayor precisión en comparación con el uso exclusivo de la demanda histórica.

# Tabla de contenido

1.	Introducción.....	1
1.1.	Antecedentes Generales.....	1
1.1.1.	Contexto sector pecuario en Chile.....	1
1.1.2.	AGROSUPER.....	2
1.2.	Justificación del proyecto.....	8
1.3.	Rol del estudiante.....	10
1.4.	Objetivos.....	11
1.5.	Alcances.....	11
2.	Marco conceptual.....	12
2.1.	Series de tiempo.....	12
2.2.	Redes neuronales.....	13
2.2.1.	Red FeedForward.....	13
2.2.2.	Redes Neuronales Recurrentes.....	15
2.2.3.	Algoritmo de optimización dentro de la red.....	19
2.3.	Entrenamiento y evaluación de redes neuronales.....	19
2.4.	Indicadores de error y métricas.....	20
2.4.1.	MAE.....	20
2.4.2.	MAPE.....	21
2.4.3.	Accuracy.....	21
2.5.	Herramientas.....	21
2.5.1.	TensorFlow.....	22
2.6.	Elasticidad.....	22
2.7.	SAP IBP.....	23
2.8.	CRISP-DM.....	23
3.	Metodología.....	24
4.	Desarrollo del proyecto.....	26
4.1.	Comprensión y preparación de los datos.....	26
4.1.1.	Base demanda y factura histórica:.....	26
4.1.2.	Base de precio histórico:.....	27
4.1.3.	Base utilizada para el procesamiento:.....	27
4.1.4.	Base demanda real:.....	28
4.2.	Variables.....	28
4.2.1.	Características de las variables de estudio.....	28
4.3.	Modelamiento.....	30

4.4. Evaluación .....	31
5. Resultados.....	32
5.1. Sector Elaborados .....	32
5.2. Sector Cecinas.....	36
6. Conclusiones.....	40
6.1. Conclusión y recomendaciones.....	40
7. Aprendizajes .....	41
8. Bibliografía.....	43
ANEXOS .....	45
Anexo A: Tablas.....	45
Sector Elaborados.....	45
Sector Cecinas .....	47
Anexo B: Gráficos.....	50
Sector Elaborados.....	50
Sector Cecinas .....	60

# Índice de Ilustraciones

Ilustración 1: Organigrama de Agrosuper .....	3
Ilustración 2: Porcentaje de ventas totales Agrosuper 2022.....	3
Ilustración 3: Organigrama área de Cadena de Suministro – Planificación de Demanda.....	6
Ilustración 4: Celda de red neuronal LSTM.....	18
Ilustración 5: Diagrama de metodología CRISP DM.....	25
Ilustración 6: Gráficos de variables en el tiempo sector Cecinas.....	29
Ilustración 7: Gráficos de variables en el tiempo sector Elaborado .....	29
Ilustración 8: Gráfico resumen sector elaborados ISOWeek 23.....	33
Ilustración 9: Gráfico promedio semanal sector elaborados .....	34
Ilustración 10: Diferencia del Accuracy de redes con variables vs caso base, sector elaborados .	35
Ilustración 11: Gráfico resumen sector cecinas ISOWeek 23 .....	37
Ilustración 12: Gráfico promedio semanal sector cecinas .....	38
Ilustración 13: Diferencia del Accuracy de redes con variables vs caso base, sector cecinas .....	39
Ilustración 14: Gráfico resumen sector elaborados ISOWeek 24.....	50
Ilustración 15: Gráfico resumen sector elaborados ISOWeek 25.....	50
Ilustración 16: Gráfico resumen sector elaborados ISOWeek 26.....	51
Ilustración 17: Gráfico caso base, sector elaborados.....	52
Ilustración 18: Gráfico red LSTM con variable precio, sector elaborados .....	53
Ilustración 19: Gráfico red LSTM con variable factura, sector elaborados .....	54
Ilustración 20: Gráfico red LSTM con variable elasticidad, sector elaborados .....	55
Ilustración 21: Gráfico red LSTM con variables precio y elasticidad, sector elaborados.....	56
Ilustración 22: Gráfico red LSTM con variables precio y factura, sector elaborados.....	57
Ilustración 23: Gráfico red LSTM con variables elasticidad y factura, sector elaborados.....	58
Ilustración 24: Gráfico red LSTM con todas las variables, sector elaborados .....	59
Ilustración 25: Diferencia del Accuracy de redes con variables vs caso base, sector elaborados .	59
Ilustración 26: Gráfico resumen sector cecinas ISOWeek 24 .....	60
Ilustración 27: Gráfico resumen sector cecinas ISOWeek 25 .....	60
Ilustración 28: Gráfico resumen sector cecinas ISOWeek 26 .....	61
Ilustración 29: Gráfico caso base, sector cecinas .....	62
Ilustración 30: Gráfico red LSTM con variable precio, sector cecinas .....	63
Ilustración 31: Gráfico red LSTM con variable factura, sector cecinas .....	64
Ilustración 32: Gráfico red LSTM con variable elasticidad, sector cecinas .....	65
Ilustración 33: Gráfico red LSTM con variables precio y elasticidad, sector cecinas .....	66
Ilustración 34: Gráfico red LSTM con variables precio y factura, sector cecinas .....	67
Ilustración 35: Gráfico red LSTM con variables elasticidad y factura, sector cecinas .....	68
Ilustración 36: Gráfico red LSTM con todas las variables, sector cecinas .....	69
Ilustración 37: Diferencia del Accuracy de redes con variables vs caso base, sector cecinas .....	69

# Índice de Tablas

Tabla 1: Distribución de variables para cada red LSTM.....	31
Tabla 2: Resultado red LSTM utilizando solo demanda histórica, sector elaborados .....	32
Tabla 3: Diferencia del Accuracy de las redes con variables vs Accuracy del caso base, sector elaborados.....	34
Tabla 4: Resultado red LSTM utilizando solo demanda histórica, sector cecinas .....	36
Tabla 5: Diferencia del Accuracy de las redes con variables vs Accuracy del caso base, sector cecinas .....	38
Tabla 6: Resultado red LSTM utilizando variable precio, sector elaborados .....	45
Tabla 7: Resultado red LSTM utilizando variable factura, sector elaborados .....	45
Tabla 8: Resultado red LSTM utilizando variable elasticidad, sector elaborados .....	46
Tabla 9: Resultado red LSTM utilizando variables precio y elasticidad, sector elaborados.....	46
Tabla 10: Resultado red LSTM utilizando variables precio y factura, sector elaborados.....	46
Tabla 11: Resultado red LSTM utilizando variables elasticidad y factura, sector elaborados.....	47
Tabla 12: Resultado red LSTM utilizando variables precio, factura y elasticidad, sector elaborados.....	47
Tabla 13: Resultado red LSTM utilizando variable precio, sector cecinas .....	47
Tabla 14: Resultado red LSTM utilizando variable factura, sector cecinas .....	48
Tabla 15: Resultado red LSTM utilizando variable elasticidad, sector cecinas .....	48
Tabla 16: Resultado red LSTM utilizando variables precio y elasticidad, sector cecinas .....	48
Tabla 17: Resultado red LSTM utilizando variables precio y factura, sector cecinas .....	49
Tabla 18: Resultado red LSTM utilizando variables elasticidad y factura, sector cecinas .....	49
Tabla 19: Resultado red LSTM utilizando variables precio, factura y elasticidad, sector cecinas	49

# **1. Introducción**

## **1.1. Antecedentes Generales**

### **1.1.1. Contexto sector pecuario en Chile**

De acuerdo con el Banco Central de Chile, el PIB del sector agroindustrial de alimentos en Chile en 2022 fue de 322.258 millones de pesos chilenos, esto representa aproximadamente el 2% del PIB total de Chile de este año. [1]. Una de las aristas de este sector es la industria cárnica, la cual se caracteriza por su diversidad y continua evolución [2], lo que ha contribuido a su fortalecimiento en los últimos años.

Este subsector se distingue por la producción de una amplia gama de carnes y productos procesados, donde la oferta cárnica en el país abarca carnes rojas, como el vacuno y el cerdo, así como carnes blancas, entre las que se encuentran el pollo y el pavo. Además, se extiende una amplia variedad de productos cárnicos procesados derivados de las carnes previamente mencionadas, entre los que se incluyen las hamburguesas y las cecinas. Tal diversificación ha facilitado la satisfacción de un amplio espectro de gustos y preferencias, tanto en el mercado nacional como en el internacional.

Para comprender mejor la relevancia de la demanda de carne a nivel nacional, es crucial considerar los siguientes aspectos sobre la industria cárnica en Chile, en el año 2022 se caracterizó por la predominancia de la producción avícola, que está altamente integrada y controlada por un número limitado de productores, alcanzando durante este año una cifra de 675 mil toneladas producidas. En paralelo, la producción de carne de cerdo, que ocupa el segundo lugar, cuenta con una producción total de 584 mil toneladas, destacando que cerca del 40% de esta producción se dirige a las exportaciones. En tercer lugar, encontramos la producción de carne bovina, que sumó un total de 200 mil toneladas [3].

Por otro lado, respecto al consumo de carne en 2021, el pollo fue la carne más consumida en el país, representando el 39% del total, seguido de la carne de res, con un 33%, y en tercer lugar la carne de cerdo, que constituye el 23% del consumo total de carne [2].

En Chile, algunas de las principales empresas dedicadas a la producción y comercialización de productos de origen animal incluyen Agrosuper, Ariztía, Coexca, Frigosorno, Productos Fernández, Multiexport Foods y Coexpan Chile. Estas empresas son reconocidas en el sector y desempeñan un papel clave en la producción y distribución de carnes de cerdo, aves, bovino, mariscos y otros productos relacionados.

## **1.1.2. AGROSUPER**

### **1.1.2.1 Contexto empresa**

El presente trabajo de título se realizó en Agrosuper S.A., la cual es una compañía de alimentos que elabora y comercializa productos de proteína animal de pollos, cerdos, pavos, salmones y procesados. El proceso productivo se encuentra integrado verticalmente y está compuesto por fábricas de alimento para los animales, granjas de crianza, plantas de procesos, centros de distribución, centros de mar y oficinas comerciales, que le permiten llegar con un amplio portafolio de productos a clientes y consumidores en Chile y el mundo, donde se le reconoce principalmente por sus marcas, las cuales son Super Pollo, La Crianza, Súper Beef, Super Cerdo, King y Sopraval.

Con respecto a su historia, Agrosuper inicia sus actividades el año 1955 con la producción de huevos en la localidad de Doñihue, VI Región y cinco años más tarde, Gonzalo Vial, su fundador, decide expandir el negocio hacia la producción y comercialización de pollos vivos, a partir de este punto la empresa ha crecido exponencialmente en los últimos años, consolidándose como una de las empresas líderes en su sector en Chile y en América Latina [4].

La misión de Agrosuper es producir y ofrecer alimentos de alta calidad, de manera sustentable y a precios competitivos, satisfaciendo las necesidades y expectativas de sus clientes y consumidores, y generando valor para sus accionistas, colaboradores y la comunidad en general [5]. Todo esto con el fin de cumplir con su propósito el cual es “Alimentar lo bueno de la vida todos los días” [4].

En el 2022, Agrosuper logró un total de ventas de \$4.179 millones de dólares. De este monto, el 60,3% correspondió a exportaciones y el 39,7% al mercado nacional [6].

### **1.1.2.2 Estructura organizacional**

En cuanto a su tamaño, la empresa emplea a más de 19.000 personas en Chile y en el extranjero [6].

Además, Agrosuper tiene una estructura organizacional bien definida, con áreas específicas para la producción, la planificación de demanda, las ventas, las finanzas, la logística, entre otras. Esto para cada uno de los tres segmentos importantes de Agrosuper, partiendo con el segmento de carnes el cual se encarga de las líneas de negocios de pollo, cerdo, pavo y procesados, luego, el corporativo, el cual se encarga de área más transversales de la compañía y, por último, el segmento acuícola, donde este se encarga de toda la línea de negocios de salmón.

Lo descrito en párrafos anteriores se puede apreciar en el siguiente organigrama de Agrosuper:

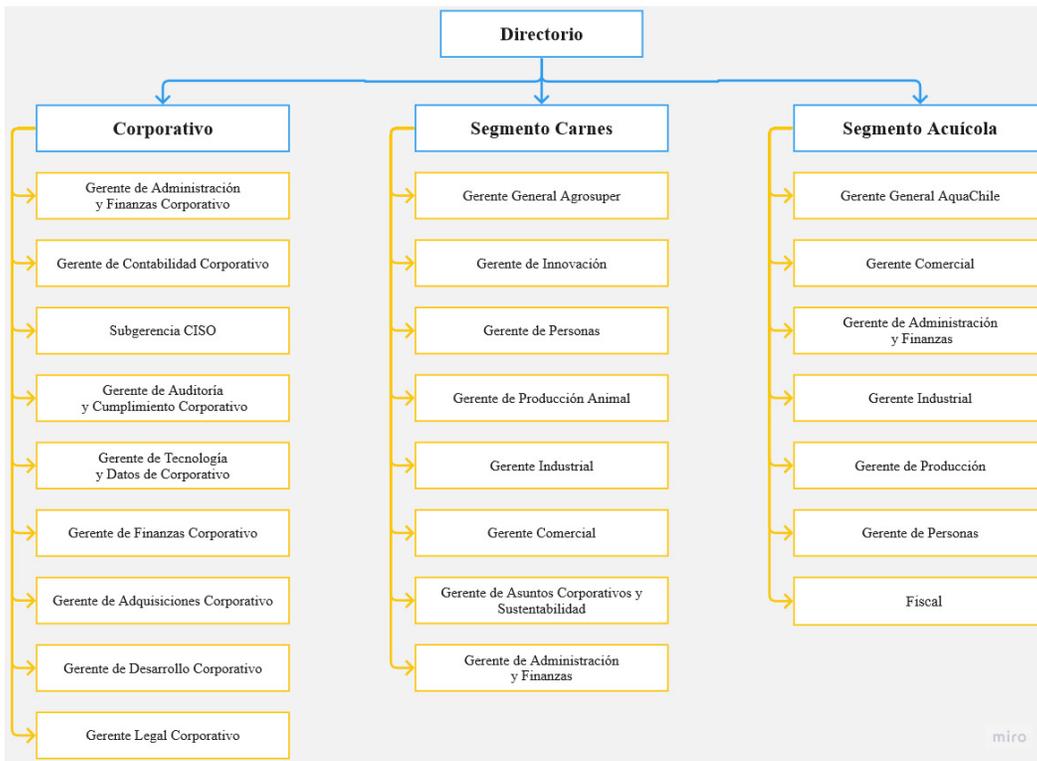


Ilustración 1: Organigrama de Agrosuper

Fuente: Adaptación de organigrama de Reporte Integrado Agrosuper 2022

### 1.1.2.3 Líneas de negocio y proceso productivo

Las principales líneas de negocio que tiene Agrosuper se representan en el siguiente gráfico, donde también se ilustra sus porcentajes correspondientes de las ventas totales del año 2022 [6].

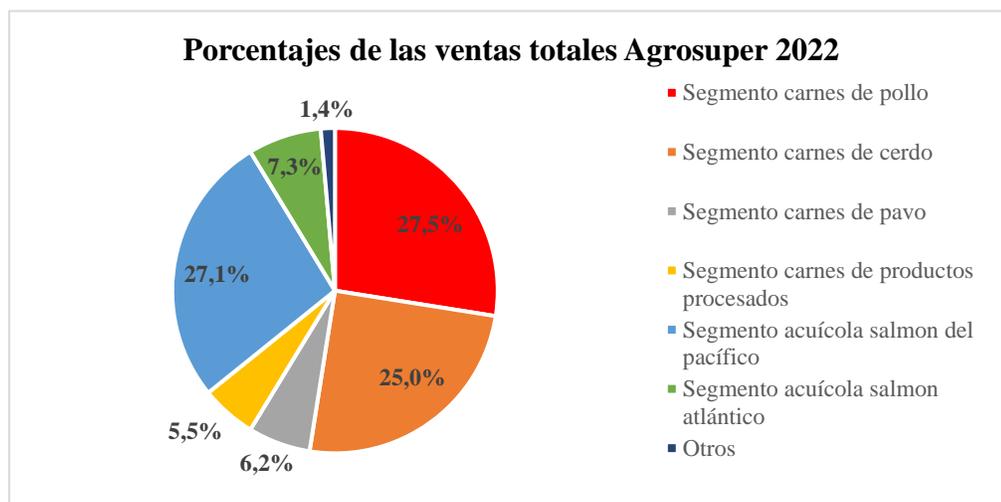


Ilustración 2: Porcentaje de ventas totales Agrosuper 2022

Fuente: Adaptación de Reporte Integrado Agrosuper 2022

El restante 1,4% de las ventas totales corresponde a la comercialización de productos de otro tipo de categoría, tales como hortalizas de la marca La Crianza.

En relación con el proceso productivo, este comienza en las plantas de elaboración de alimentos para animales. Luego, la materia prima se transporta a los sectores de crianza, donde los animales son criados para posteriormente ser llevados a las plantas de procesamiento. Tras este paso, la producción se divide en dos líneas de distribución: una para sus clientes nacionales y otra para sus clientes internacionales. Finalmente, los productos son comercializados y llegan a sus respectivos consumidores [6].

#### **1.1.2.4 Sector de productos procesados**

El presente trabajo de título se enfoca principalmente en el sector de productos procesados, donde en el año 2022 las ventas totales fueron de \$229 millones de dólares, siendo el 99,2% generado por ventas en el mercado nacional y el 0,8% por exportaciones [6]. Esta línea de negocio se divide en en dos categorías: productos de cecinas y productos elaborados.

Los principales clientes de Agrosuper del sector de productos procesados se distribuyen en cuatro canales, los cuales son:

1. Supermercados: Se refiere a los establecimientos de venta al por menor que ofrecen una amplia gama de productos alimenticios y no alimenticios, tales como Jumbo, Unimarc, Santa Isabel, entre otros.
2. Canal Tradicional: Engloba diversos puntos de venta a nivel nacional, como minimarkets y carnicerías, entre otros.
3. Foodservice: Este canal está dirigido a clientes que operan en el sector de servicios de alimentos, como casinos, instituciones y restaurantes de comida rápida.
4. Grandes Clientes: Este canal se refiere a pedidos de particulares, que también pueden incluir ventas a otras empresas.

En cuanto a los productos específicos, en el caso de las cecinas, se destacan jamones, salchichas y pechuga de pavo, mientras que en la categoría de productos elaborados se incluyen hamburguesas, empanizados como nuggets de pollo y carne molida de otras líneas de negocio.

Por otro lado, la línea de negocio de productos procesados de Agrosuper tuvo una participación en el mercado de aproximadamente un 7% durante los meses de diciembre, enero y febrero de 2022-2023. Esta cifra se sitúa por debajo de sus principales competidores, como CIAL Alimentos, con un 42%, y Productos Fernández, con un 32%.<sup>1</sup>

---

<sup>1</sup> Indicador Nielsen, información interna de Agrosuper.

Aproximadamente, se estima una demanda periódica semanal de 1.100 toneladas en este sector de productos procesados, siendo 600 toneladas correspondientes a productos elaborados y 500 toneladas a productos de cecinas, donde los principales clientes se distribuyen en cuatro canales:

### **1.1.2.5 Área de cadena de suministro**

El área de cadena de suministro es una de las áreas que depende bajo la responsabilidad de la Gerencia Comercial del Segmento Carnes de Agrosuper. Es una unidad organizativa encargada de gestionar y coordinar todas las actividades relacionadas con el flujo de productos, desde los proveedores hasta los clientes. Su principal rol es asegurar que los productos sean entregados de manera eficiente, oportuna y con la calidad requerida, optimizando los recursos disponibles y minimizando los costos.

El área de cadena de suministro de Agrosuper tiene como objetivo principal planificar, coordinar y controlar las diferentes etapas del proceso de suministro. Esto incluye la gestión de la adquisición de materias primas, la planificación de la producción, el control de inventarios, la distribución y la logística, así como la coordinación con proveedores y clientes.

Su principal rol implica:

1. Planificación de la demanda: Analizar y predecir la demanda de los productos para garantizar la disponibilidad adecuada de inventario y evitar excesos o faltantes.
2. Gestión de proveedores: Establecer relaciones sólidas con los proveedores, negociar contratos y acuerdos, y asegurarse de que se cumplan los estándares de calidad y plazos de entrega.
3. Gestión de inventarios: Controlar los niveles de inventario de manera eficiente para evitar sobrecostos o escasez de productos, utilizando técnicas de gestión y sistemas de información.
4. Planificación de la producción: Establecer planes de producción óptimos y coordinar con los departamentos internos para asegurar la producción en tiempo y forma.
5. Distribución y logística: Coordinar la distribución y el transporte de los productos a los clientes, optimizando rutas, reduciendo costos y asegurando la entrega puntual.

### **1.1.2.6 Área de planificación de demanda**

El trabajo de título se realizó en el área de Planificación de Demanda, una de las secciones bajo la responsabilidad de la Gerencia de Cadena de Suministro, ubicada en la matriz central en la comuna de Rancagua. El objetivo principal del área de planificación de demanda en Agrosuper es anticipar, estimar y administrar de manera precisa la demanda futura para los productos que

ofrece la empresa. Esta planificación de demanda forma la base para decisiones estratégicas que influyen en aspectos como la oferta, la capacidad de producción y los niveles de inventario.

La principal responsabilidad de esta área es proyectar la demanda futura en cada periodo correspondiente, lo que a su vez establece las bases para la planificación operativa y táctica. Este proceso de planificación es esencial ya que proporciona una comprensión clara de cómo evoluciona el negocio en términos de demanda y cuál sería la planificación de esta en cada periodo.

En consecuencia, el rol de la planificación de demanda en la empresa es vital. Al proporcionar claridad sobre la evolución de la demanda, guía la toma de decisiones en todas las otras áreas de la empresa, desde la producción hasta las ventas y la gestión de inventario.

De lo anterior, se ilustra el siguiente organigrama del área de Cadena de Suministros, en conjunto a lo que conforma el área de Planificación de Demanda:

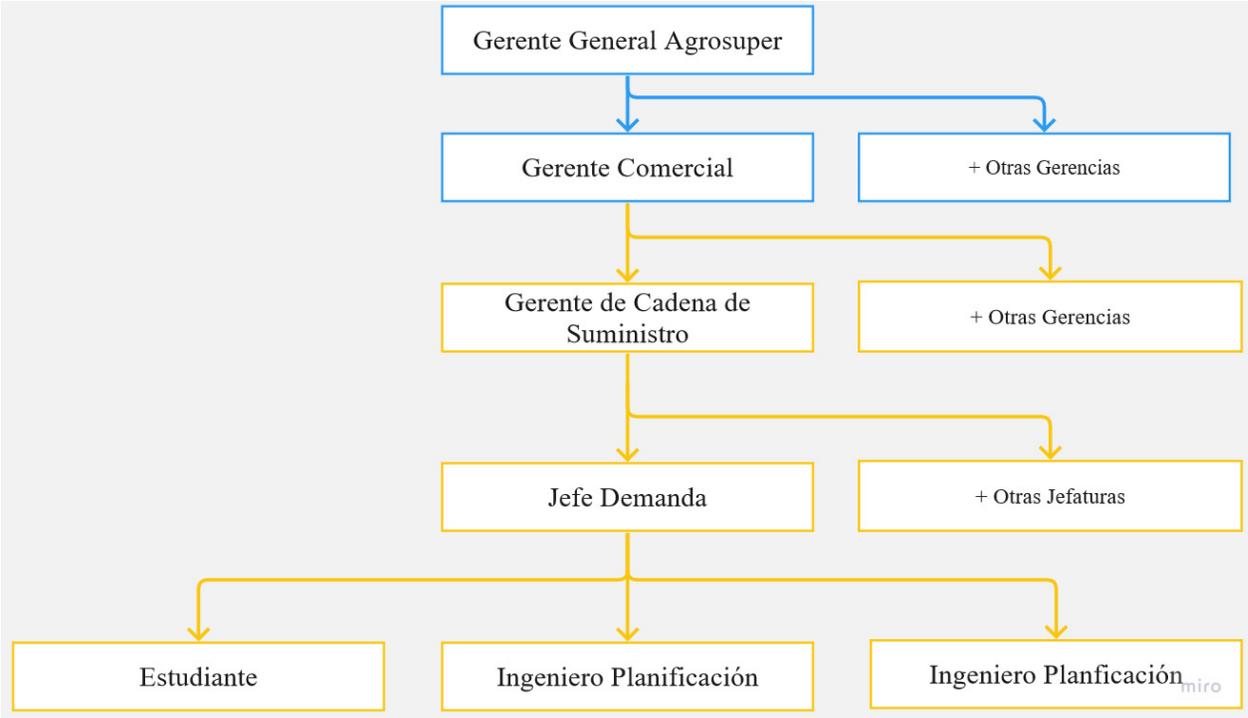


Ilustración 3: Organigrama área de Cadena de Suministro – Planificación de Demanda  
Fuente: Adaptación de organigrama de Mi Mundo Super – Página laboral de Agrosuper

A continuación, se detallan algunas de las responsabilidades y funciones clave del rol de la planificación de la demanda:

1. Recopilación y análisis de datos: En un inicio se recopila información relevante, como históricos de ventas, datos del mercado, tendencias, promociones y otros factores que puedan influir en la demanda. Luego, se analiza estos datos para identificar patrones, estacionalidades, cambios en la demanda y otros factores importantes.
2. Pronóstico de la demanda: Utilizando métodos y técnicas estadísticas, tales como la utilización de modelos de predicción, se generan pronósticos de la demanda futura de los productos que ofrece Agrosuper para cada línea de negocio. Esto implica considerar diferentes variables, como la demanda histórica, la evolución del mercado, eventos especiales, estacionalidad, entre otros.
3. Colaboración con otras áreas: Planificación de demanda trabaja en estrecha colaboración con otras áreas dentro de la Cadena de Suministro para recopilar información adicional, validar pronósticos y comprender mejor las necesidades y expectativas del mercado.
4. Ajuste de pronósticos: A medida que se obtienen nuevos datos o se producen cambios en las condiciones del mercado, se realiza ajustes en los pronósticos. Esto implica analizar las desviaciones entre los pronósticos previos y los resultados reales, identificar las causas de las variaciones y actualizar los pronósticos en consecuencia.
5. Recopilación de antecedentes estratégicos: La planificación de demanda desempeña un papel activo en las reuniones de los distintos departamentos y jefaturas que conforman el área de la Cadena de Suministro. En estas reuniones, se presentan los pronósticos de demanda, se comparte información sobre tendencias y factores relevantes, y se colabora con estos equipos para tomar decisiones estratégicas en términos de producción, inventario, capacidad y otras áreas relacionadas.

Uno de los desafíos del área de planificación de demanda es la mejora continua de los modelos de predicción que están en uso. En su estado actual, estos modelos se basan exclusivamente en información histórica de demanda, desglosada por los siguientes atributos:

1. Sector: Línea de negocio específico, tales como pollo, cerdo, cecina, elaborado, etc.
2. Tipo de cliente: Supermercado, Canal Tradicional, FoodService, etc.

3. Segmentación de tipo de cliente: Cliente específico como por ejemplo Supermercado Jumbo, FoodService casino, Canal Tradicional Cobertura Grande, etc.
4. Tipo de producto: Línea de producto específico, tales como salchichas, pechugas, empanizados, etc.
5. SKU (Stock Keeping Unit): Código unitario por producto específico.

Complementándose con ciertos criterios temporales de estacionalidad y aplicando el tratamiento de los datos con modelos de series de tiempo.

Se debe resaltar que el área utiliza dos métricas para medir la precisión de sus pronósticos de demanda: el MAPE y el Accuracy.

El MAPE, o Error Porcentual Absoluto Medio (Mean Absolute Percentage Error, por sus siglas en inglés), es una medida que expresa la precisión de un pronóstico en términos porcentuales. Es una métrica comúnmente utilizada en la planificación de la demanda para comparar la precisión de diferentes modelos de pronóstico.

Por otro lado, el Accuracy es el valor complementario al MAPE. Si el MAPE mide el error, el Accuracy mide el grado de exactitud de los pronósticos. Es decir, un Accuracy del 100% significaría un pronóstico perfecto, mientras que un Accuracy del 90% indicaría que los pronósticos están correctos el 90% del tiempo.

Al implementar un modelo de predicción mejorado, es posible que se observe una reducción en el valor del MAPE y un aumento correspondiente en el Accuracy. Esto implicaría un aumento en la precisión de los pronósticos de demanda.

## **1.2. Justificación del proyecto**

Dado que los modelos que utiliza el área de planificación de demanda se basan exclusivamente en información histórica de demanda, se identifica una oportunidad de mejora en la exploración de nuevas variables que podrían mejorar la precisión de la predicción de demanda. Ante esta oportunidad, el área tiene previsto implementar un proyecto piloto centrado exclusivamente en el sector de productos procesados. Esta elección radica en que, en comparación con otros sectores cárnicos, se trata de una línea de productos secundaria que, además, cuenta con una demanda estable. Por lo tanto, la incorporación y experimentación con nuevas variables no debería representar complicaciones significativas.

Esta estrategia permitiría la generalización y aplicación de las lecciones aprendidas y las mejores prácticas identificadas en este piloto, optimizando así la precisión en la predicción de la demanda a lo largo de diversas líneas de productos.

En este contexto, se presenta una oportunidad para el estudiante de asumir este desafío, aplicando los conocimientos adquiridos durante su formación académica, así como sus habilidades de comunicación para transmitir efectivamente los resultados del proyecto.

En un inicio, el estudiante tuvo la intención de utilizar el conglomerado de modelos empleados en el área de planificación de la demanda. Sin embargo, estos no pueden ser modificados, ya que están implementados en una plataforma virtual el cual tiene permisos limitados para estudiantes en práctica dentro de la empresa.

Luego, el estudiante intentó utilizar de forma manual un conglomerado de modelos de series de tiempo. Sin embargo, se enfrentó dificultades al intentar modelar y capturar un valor similar a los de la demanda de productos procesados.

Finalmente, se solicitó ayuda al área de Ciencias de Datos de Agrosuper, donde este departamento ya había probado un piloto de uso de redes neuronales para la línea de negocios de pollo.

Dado esto el estudiante propone desarrollar un nuevo modelo utilizando redes neuronales que incorpore estas nuevas variables a considerar, inspirado a lo realizado por el área de Ciencias de Datos.

El cumplimiento de este desafío traería varios beneficios para el área de Cadena de Suministro [7], tales como:

- Mejora en la precisión de las predicciones: Incorporar nuevas variables y aplicar técnicas de aprendizaje automático avanzadas, como las redes neuronales, pueden mejorar la precisión de las predicciones de demanda, lo que a su vez permite una planificación más efectiva y eficiente.
- Optimización del inventario: Una predicción más precisa de la demanda puede resultar en una mejor gestión del inventario, minimizando el exceso de stock y reduciendo la probabilidad de escasez de productos.
- Mejora de la eficiencia operativa: Con una planificación de demanda más precisa, se pueden optimizar las operaciones de producción, distribución y logística, lo que puede llevar a ahorros significativos en tiempo y costos.

La oportunidad de mejora a la que se dirige este trabajo de título radica en evaluar la precisión del pronóstico de demanda al incorporar estas nuevas variables. Para ello, el estudiante propone el desarrollo de un nuevo modelo utilizando redes neuronales que integre dichas variables, manteniendo al mismo tiempo el análisis de la estacionalidad de la demanda y el modelado de series de tiempo. En este proceso, se tomará en consideración exclusivamente el sector de productos procesados, es decir, de productos elaborados y cecinas.

Aplicar esta solución presenta múltiples ventajas:

En primer lugar, el uso de redes neuronales permite un análisis más complejo y matizado de los datos, ya que estas pueden detectar y aprender patrones no lineales y relaciones intrincadas entre las variables, por consiguiente, esta capacidad puede dar lugar a pronósticos más precisos [8].

En segundo lugar, mantener el análisis de la estacionalidad de la demanda y el modelado de series de tiempo garantiza que el modelo sigue considerando los factores temporales en la predicción de la demanda [8].

Finalmente, limitar el enfoque inicial al sector de productos procesados proporciona un contexto más controlado para probar y perfeccionar el modelo. Debido a la estabilidad de la demanda en este sector, se reduce el riesgo de factores externos imprevistos que podrían alterar los resultados. Esto permite un análisis más claro de la efectividad de las nuevas variables y las redes neuronales en la mejora de la precisión del pronóstico.

### **1.3. Rol del estudiante**

En este proyecto, el estudiante se desempeña como el principal investigador y desarrollador de la solución propuesta, con el rol de llevar a cabo la oportunidad de mejora identificada. Este rol es necesario debido a la especialización y las habilidades requeridas para el manejo y la implementación de redes neuronales, así como para el análisis de datos.

Una de las tareas principales del estudiante es identificar nuevas variables que puedan mejorar la precisión de la predicción de demanda. Esto implica realizar un análisis exploratorio sobre las bases de datos recopilados dentro de la organización.

También el estudiante es responsable del desarrollo de la red neuronal, la cual requiere habilidades avanzadas en programación y una comprensión profunda de las redes neuronales y su funcionamiento, así como de los métodos de entrenamiento y ajuste para garantizar que la red pueda manejar las variables identificadas y producir resultados precisos.

Además, el estudiante llevará a cabo la evaluación y análisis de los resultados generados por la red neuronal. En particular, se realizará un análisis comparativo utilizando un caso base que solo utiliza la demanda histórica y otras variantes de la misma red con diferentes combinaciones de las variables identificadas. De esta manera, se medirá la mejora en la precisión y eficacia de los pronósticos en cada escenario, utilizando las mismas métricas para medir la precisión de los pronósticos que utiliza el área de planificación de demanda.

Finalmente, el rol del estudiante es necesario para este proyecto, ya que requiere una combinación de habilidades analíticas y técnicas, una comprensión profunda del negocio y la capacidad para comunicar eficazmente los hallazgos a la respectiva área de trabajo, así como a otras áreas interesadas dentro de la empresa.

## **1.4. Objetivos**

### **1.4.1. Objetivo general**

Evaluar la incorporación de nuevas variables en la predicción de la demanda en el sector de productos procesados de Agrosuper, con el fin de mejorar su precisión y proporcionar evidencia empírica que respalde la eficacia de la implementación de estas variables.

### **1.4.2. Objetivos específicos**

1. Definir qué variables serán incorporadas en el modelo de predicción de demanda, para analizar su efecto en las predicciones.
2. Desarrollar una red neuronal referencial basada únicamente en la demanda histórica, además de elaborar variaciones de dicha red que incorporan las nuevas variables, con el fin de establecer un punto de comparación con el modelo inicial.
3. Evaluar la precisión predictiva de las redes desarrolladas utilizando los indicadores estadísticos MAPE y Accuracy, para identificar si hay diferencias con el caso base.

## **1.5. Alcances**

Los alcances del trabajo a realizar se pueden segmentar en varios aspectos, lo que permitirá abordar la oportunidad de mejora de manera más dinámica sin desviar la atención del proceso principal.

En primer lugar, se debe tener en cuenta el alcance del trabajo en términos de tiempo, pues debido a lo acotado que es éste. La recopilación y limpieza de los datos, la identificación de las variables, el desarrollo de la red neuronal, la evaluación del funcionamiento de la red, el análisis de los resultados y la implementación de este mismo, comprenden por un periodo mayor a un semestre académico. Por lo tanto, la fase de implementación no se considerará en este proyecto.

Cabe mencionar que inicialmente se tuvo la iniciativa de realizar el trabajo a nivel micro, pero debido a restricciones de tiempo, no es posible llevarlo a cabo de manera completa. Por esta razón, se optó por abordar el estudio a nivel macro, centrándose en la demanda del sector y las variables de interés.

En segundo lugar, el proyecto se lleva a cabo utilizando una base de datos que comprende el periodo desde mayo del año 2020 hasta mayo del año 2023. La razón de esto es que, a partir de 2020, la variable de precio se considera de manera específica por tipo de cliente, segmentación de

cliente, tipo de producto y código SKU, mientras que anteriormente se utilizaba un promedio consolidado. Debido a esto, se está omitiendo una cantidad de información histórica, lo que podría impactar en la capacidad de generar pronósticos de demanda más precisos. Esto es particularmente relevante ya que, al manejar una mayor cantidad de datos, la red neuronal se beneficia al tener más historia de entrada para su aprendizaje.

En tercer lugar, solo se están considerando variables internas de la organización. Por lo tanto, no se toman en cuenta variables externas como la competencia en determinados productos, a pesar de que estas podrían influir en el valor de la demanda real en las fechas para las que se desean hacer pronósticos y evaluar su precisión [9].

Finalmente, hay que señalar que los datos de algunos productos podrían ya no existir en la actualidad. Esto puede ser debido a que se dejó de producir, hubo un cambio de código y/o el producto fue trasladado a otra línea de negocio. Estos factores también podrían generar un error a nivel macro en la predicción de demanda.

## **2. Marco conceptual**

Dado el contexto y los objetivos de este trabajo de título, se tiene que este se sitúa dentro del ámbito de la Ciencia de Datos, una disciplina dentro de la Ingeniería Civil Industrial que busca obtener conocimientos y perspectivas a partir de datos estructurados y no estructurados a través de diversas técnicas y métodos. En particular, este trabajo de título se centra en el análisis de series de tiempo, que se define como el esfuerzo por extraer información resumida y estadística de puntos dispuestos en orden cronológico [10], correspondiendo en este caso, la demanda semanal de productos procesados en Agrosuper y el impacto de las variables a incorporar.

### **2.1. Series de tiempo**

Gracias al autor, George Box quien, con la publicación de su libro "Time Series Analysis: Forecasting and Control" en 1970, impulsó significativamente el campo de estudio de las series de tiempo, introduciendo el reconocido método Box-Jenkins que se utiliza incluso en la actualidad. [10].

Las series de tiempo son una serie de puntos de datos recopilados en intervalos regulares y ordenados en el tiempo. Con el avance de las tecnologías de programación y computación, el análisis de series de tiempo cobró mayor relevancia, especialmente desde una perspectiva estadística que se centra en la aplicación de modelos autorregresivos, debido a la correlación temporal que presentan los datos

En el contexto de predicción de la demanda, las series de tiempo se utilizan comúnmente para observar los patrones de demanda históricos y usarlos para prever la demanda futura. Algunos de los patrones que se pueden identificar a partir de las series de tiempo incluyen:

1. **Tendencia:** Un patrón a largo plazo en los datos, como por ejemplo un aumento general en la demanda a lo largo del tiempo.
2. **Estacionalidad:** Patrones que se repiten a intervalos regulares, como por ejemplo un aumento en la demanda de salchichas parrilleras durante las fiestas patrias.
3. **Ciclicidad:** Patrones que ocurren a intervalos irregulares.
4. **Irregularidad o ruido:** La variabilidad en los datos que no puede explicarse por los patrones anteriores.

Actualmente, los modelos de Machine Learning lideran los estudios de datos, siendo el modelo presentado en el artículo "Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting" [11] el estado del arte para las series de tiempo. Este se basa principalmente en el uso de redes neuronales, que identifican patrones en los datos y los replican en las predicciones.

## **2.2. Redes neuronales**

Como contexto, las redes neuronales son un modelo de computación inspirado en el funcionamiento del cerebro humano [12], específicamente en cómo las neuronas se conectan y transmiten información entre sí. Estos sistemas aprenden a realizar tareas al considerar ejemplos, generalmente sin ser programados con reglas específicas para la tarea.

Las redes neuronales son capaces de identificar patrones complejos, aprender de los datos y hacer predicciones, lo que las hace útiles para tareas como clasificación, regresión, agrupación y detección de anomalías.

A continuación, se presentarán 2 modelos de redes neuronales, entre los cuales destaca la Feedforward. Este modelo ha sido utilizado como una iniciativa piloto por el equipo de ciencia de datos de Agrosuper y ha servido como fuente de inspiración para este trabajo de título.

### **2.2.1. Red FeedForward**

La red neuronal Feedforward es uno de los tipos más simples de redes neuronales. En este modelo, la información se mueve en una sola dirección, desde la entrada hasta la salida, sin formar ciclos ni bucles. Esto significa que las capas de neuronas están completamente conectadas y que la información siempre avanza, nunca se retroalimenta hacia atrás. Son especialmente útiles en tareas de clasificación y regresión [13].

Para explicar cómo funciona si tiene lo siguiente:

Supongamos que tenemos una red feedforward con  $L$  capas, donde la capa 0 es la capa de entrada y la capa  $L$  es la capa de salida.

### 1. Entrada de datos:

El vector de entrada en la capa 0 se denota como  $x \in \mathbb{R}^n$ , donde  $n$  es el número de características de entrada.

### 2. Capas ocultas:

Cada capa oculta  $l \in \{1, 2, \dots, L - 1\}$  consta de  $m_l$  neuronas. Sea  $a_l \in \mathbb{R}^{m_l}$  el vector de activación de la capa  $l$  y  $a_{lj}$  la activación de la  $j$ -ésima neurona en la capa  $l$ .

### 3. Pesos y sesgos:

Los pesos de las conexiones entre las neuronas de la capa  $l$  y la capa  $l + 1$  se representan por  $W_l \in \mathbb{R}^{m_{l+1} \times m_l}$ , donde la fila  $j$ -ésima de  $W_l$  denota los pesos de la  $j$ -ésima neurona en la capa  $l + 1$  conectada a las neuronas en la capa  $l$ . Los sesgos de la capa  $l + 1$  se denotan por  $b_{l+1} \in \mathbb{R}^{m_{l+1}}$ .

### 4. Función de activación:

La función de activación de las neuronas en la capa  $l$  se denota como  $f_l: \mathbb{R} \rightarrow \mathbb{R}$ . La función de activación en este caso es la función sigmoide denotada de la siguiente manera:

$$f(x) = \frac{1}{1 + e^{-x}}$$

Donde "e" es la base del logaritmo natural

### 5. Propagación hacia adelante:

Para cada capa  $l \in \{1, 2, \dots, L - 1\}$ , la activación  $a_l$  se calcula de la siguiente manera:

$$a_l = f_l(W_l * a_{l-1} + b_l)$$

Donde "\*" denota la multiplicación matricial.

### 6. Capa de salida:

La capa  $L$  (capa de salida) tiene  $k$  neuronas de salida, donde  $k$  es el número de clases o valores de salida. La activación de la neurona  $j$  en la capa  $L$  se denota como  $a_{Lj}$ .

### 7. Cálculo del error:

Dado un conjunto de entrenamiento con etiquetas de salida  $y_{train} \in \mathbb{R}^k$ , el error se puede calcular utilizando una métrica, en este caso se utilizó el error absoluto medio (MAE):

$$error = MAE(a_L, y_{train})$$

### **8. Retropropagación del error:**

El objetivo es minimizar el error ajustando los pesos y sesgos de la red. Para hacer esto, se utiliza el algoritmo de retropropagación del error para calcular el gradiente del error con respecto a los pesos y sesgos de la red.

### **9. Actualización de pesos y sesgos:**

Utilizando el gradiente del error, los pesos y sesgos se actualizan iterativamente utilizando un algoritmo de optimización, como el descenso del gradiente. La regla general para la actualización de los pesos y sesgos es:

$$W_l = W_l - \eta * \frac{\partial error}{\partial W_l}$$
$$b_l = b_l - \eta * \frac{\partial error}{\partial b_l}$$

Donde  $\eta$  es la tasa de aprendizaje que controla la velocidad de convergencia del algoritmo de optimización, el algoritmo utilizado es Adam (Adaptive Moment Estimation), este será explicado más adelante en este informe.

### **10. Repetir el proceso:**

Los pasos 5 a 9 se repiten iterativamente hasta que se alcance un criterio de convergencia, como el número máximo de iteraciones.

En resumen, la propagación hacia adelante de una red feedforward implica calcular las activaciones de las neuronas en cada capa utilizando pesos y sesgos, aplicando una función de activación no lineal. Luego, se calcula el error en la capa de salida y se utiliza la retropropagación del error para ajustar los pesos y sesgos, con el objetivo de minimizar el error a medida que se actualizan iterativamente.

## **2.2.2. Redes Neuronales Recurrentes**

Las redes neuronales recurrentes (RNR) son un tipo de arquitectura de redes neuronales diseñadas para trabajar con datos secuenciales o datos que tienen una estructura temporal [14].

La principal diferencia entre las RNR y las redes feedforward es la capacidad de las RNR para utilizar información anterior y tener conexiones recurrentes, lo que les permite capturar relaciones temporales en los datos. Mientras que en una red feedforward la información fluye en una dirección, de la capa de entrada a la capa de salida sin retroalimentación, en una RNR la información fluye en bucles, donde la salida anterior se utiliza como entrada en el siguiente paso de tiempo.

Dentro de las RNR, existen variantes que pueden abordar el desafío de capturar dependencias a largo plazo en secuencias, una de esas variantes es la red neuronal llamada Long Short-Term Memory (LSTM).

## 1.2.1 Long Short-Term Memory

De acuerdo a los autores Haşim Sak, Andrew Senior y Françoise Beaufays en su publicación “Long Short-Term Memory Recurrent Neural Network Architectures”, la red neuronal LSTM es una variante de las redes neuronales recurrentes, diseñada para abordar el problema de la desaparición del gradiente en el entrenamiento de redes recurrentes tradicionales. La LSTM utiliza unidades de memoria y puertas internas para capturar y recordar información relevante a largo plazo en secuencias de datos [14].

Para explicar cómo funciona se tiene lo siguiente:

Supongamos que tenemos una red LSTM con  $L$  capas, donde la capa 0 es la capa de entrada y la capa  $L$  es la capa de salida.

### 1. Entrada de datos:

El vector de entrada en la capa 0 se denota como  $x \in \mathbb{R}^n$ , donde  $n$  es el número de características de entrada.

### 2. Capas LSTM:

Cada capa LSTM  $l \in \{1, 2, \dots, L\}$  consta de  $m_l$  celdas LSTM. Sea  $h_l \in \mathbb{R}^{m_l}$  el vector de estado oculto de la capa  $l$  y  $c_l \in \mathbb{R}^{m_l}$  el vector de estado de la celda de memoria de la capa  $l$ .

### 3. Pesos y sesgos:

Los pesos de las conexiones entre las celdas LSTM de la capa  $l$  y la capa  $l + 1$  se representan por  $W_l \in \mathbb{R}^{\{m_{l+1} \times m_l\}}$ , donde la fila  $j$ -ésima de  $W_l$  denota los pesos de la  $j$ -ésima celda LSTM en la capa  $l + 1$  conectada a las celdas LSTM en la capa  $l$ . Los sesgos de la capa  $l + 1$  se denotan por  $b_{l+1} \in \mathbb{R}^{\{m_{l+1}\}}$ .

#### 4. Propagación hacia adelante:

Para cada capa  $l \in \{1, 2, \dots, L\}$ , la propagación hacia adelante en una celda LSTM se realiza en tres pasos: olvido, entrada y salida.

1. Olvido:

$$f_l = \sigma(W_{lf} * h_{l-1} + U_{lf} * x + b_{lf})$$

2. Entrada:

$$i_l = \sigma(W_{li} * h_{l-1} + U_{li} * x + b_{li})$$
$$c_l = \tanh(W_{lc} * h_{l-1} + U_{lc} * x + b_{lc})$$

3. Salida:

$$o_l = \sigma(W_{lo} * h_{l-1} + U_{lo} * x + b_{lo})$$
$$h_l = o_l \odot \tanh(c_l)$$

Donde " $\odot$ " denota la multiplicación elemento por elemento,  $\sigma$  es la función sigmoide y  $\tanh$  la función tangente hiperbólica.

#### 5. Capa de salida:

La capa  $L$  (capa de salida) tiene  $k$  neuronas de salida, donde  $k$  es el número de valores de salida. La activación de la neurona  $j$  en la capa  $L$  se denota como  $a_{Lj}$ .

#### 6. Cálculo del error:

Dado un conjunto de entrenamiento con etiquetas de salida  $y_{train} \in \mathbb{R}^k$ , el error se puede calcular utilizando una métrica, en este caso se utilizó el error absoluto medio (MAE):

$$error = MAE(a_L, y_{train})$$

#### 7. Retropropagación del error:

Al igual que en una red feedforward estándar, el error se retropropaga desde la capa de salida hacia las capas anteriores para calcular el gradiente del error con respecto a los pesos y sesgos de la red LSTM.

#### 8. Actualización de pesos y sesgos:

Utilizando el gradiente del error, los pesos y sesgos se actualizan iterativamente utilizando un algoritmo de optimización (Adam), como el descenso del gradiente. La regla general para la actualización de los pesos y sesgos es similar a la de una red feedforward:

$$W_l = W_l - \eta * \frac{\partial error}{\partial W_l}$$

$$U_l = U_l - \eta * \frac{\partial error}{\partial U_l}$$

$$b_l = b_l - \eta * \frac{\partial error}{\partial b_l}$$

Donde  $\eta$  es la tasa de aprendizaje.

### 9. Repetir el proceso:

Los pasos 4 a 8 se repiten iterativamente hasta que se alcance un criterio de convergencia, como el número máximo de iteraciones.

En resumen, una red LSTM utiliza celdas LSTM para capturar y recordar información a largo plazo a través del tiempo. La propagación hacia adelante en una celda LSTM implica pasos de olvido, entrada y salida para actualizar el estado de la celda y el estado oculto. La retropropagación del error y la actualización de pesos y sesgos se realizan de manera similar a una red feedforward estándar.

En la siguiente ilustración se tiene un ejemplo de una celda de la red neuronal LSTM con los parámetros mencionados:

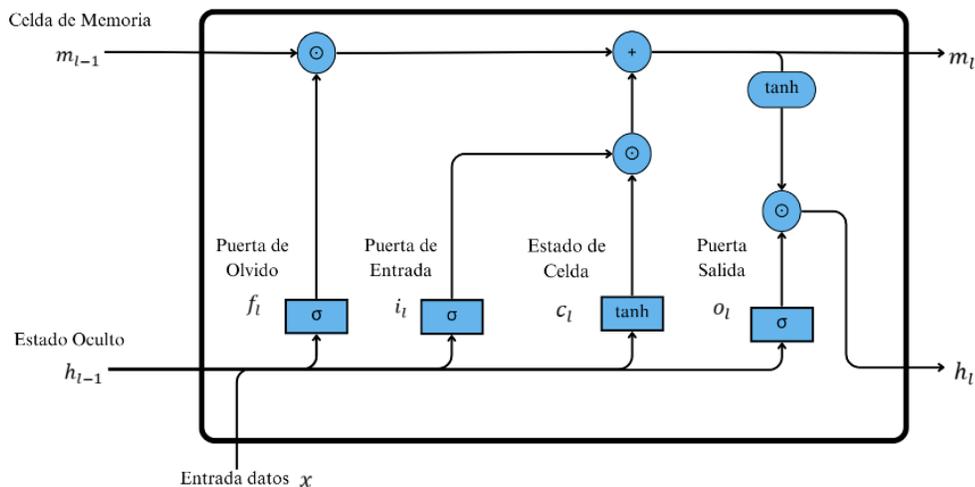


Ilustración 4: Celda de red neuronal LSTM

Fuente: Elaboración propia del estudiante utilizando Canva

### **2.2.3. Algoritmo de optimización dentro de la red**

Adam (Adaptive Moment Estimation) es un algoritmo de optimización ampliamente utilizado en el entrenamiento de redes neuronales y otros modelos de aprendizaje automático. Fue propuesto por Diederik P. Kingma y Jimmy Ba en 2015 [15].

El algoritmo de optimización Adam utiliza dos conceptos clave: momento y adaptación de la tasa de aprendizaje. El momento permite al algoritmo acumular una media móvil de los gradientes pasados, lo que ayuda a acelerar la convergencia y evitar los mínimos locales. La adaptación de la tasa de aprendizaje ajusta la tasa de aprendizaje para cada parámetro de la red de acuerdo con su importancia relativa.

Teniendo esto, el algoritmo se adapta automáticamente con tasas de aprendizaje para cada parámetro de la red neuronal en función de sus estimaciones de primer y segundo momento de los gradientes.

## **2.3. Entrenamiento y evaluación de redes neuronales**

Una vez seleccionada la red neuronal a utilizar, es necesario ajustarla a los datos que serán utilizados y evaluar su precisión en la predicción de la demanda. En este sentido, se deben considerar tres aspectos fundamentales:

1. Conjunto de entrenamiento y prueba del modelo: Para construir estos conjuntos, las observaciones se dividen generalmente de manera aleatoria en dos partes, con el objetivo de evitar sesgos en la red. Sin embargo, debido a la naturaleza temporal de los datos, es importante mantener el orden temporal en la información. Una de las particiones se utiliza para entrenar la red neuronal y realizar su ajuste, mientras que la otra se utiliza para validar la red ya entrenada y estimar su error de generalización. El objetivo es minimizar el error y evitar el sobreajuste de los datos.
2. Error: Este se refiere a la discrepancia entre los datos reales y los datos predichos por la red en el conjunto de validación. Se busca minimizar este error para lograr una mayor precisión en las predicciones.
3. Métricas de desempeño: Son medidas utilizadas para evaluar la calidad de la red y el error de este. En este caso, se utilizarán las métricas MAPE (Error Porcentual Absoluto Medio) y Accuracy para evaluar el desempeño de la red.

En el contexto de las redes LSTM (Long Short-Term Memory), el número de épocas y el tamaño del batch son dos conceptos relacionados con el entrenamiento de la red neuronal.

El número de épocas se refiere al número de veces que el modelo de LSTM recorre el conjunto de datos completo durante el entrenamiento. Cada época consiste en un ciclo completo de pasar todos los ejemplos de entrenamiento por la red y ajustar los pesos para minimizar la función de pérdida. En cada época, el modelo actualiza sus parámetros basándose en los errores cometidos durante el proceso de propagación hacia atrás (backpropagation).

El tamaño del batch (también conocido como tamaño del lote o batch size) se refiere a la cantidad de ejemplos de entrenamiento que se utilizan en una sola iteración del algoritmo de entrenamiento antes de que se realice una actualización de los pesos del modelo. En lugar de actualizar los pesos después de cada ejemplo de entrenamiento individualmente, es más común agrupar varios ejemplos en un lote y calcular el promedio de los errores de predicción en ese lote. Luego, se realiza una actualización de los pesos en función de ese promedio.

El tamaño del batch tiene un impacto en la velocidad y la estabilidad del entrenamiento de una red LSTM. Algunos de los efectos de diferentes tamaños de batch son los siguientes:

- **Tamaño de batch grande:** puede acelerar el proceso de entrenamiento al procesar más ejemplos simultáneamente. Sin embargo, requiere más memoria, y los gradientes se calculan en función de un promedio más grande, lo que puede hacer que la convergencia sea más difícil de alcanzar.
- **Tamaño de batch pequeño:** puede requerir más épocas para converger, ya que los pesos se actualizan con mayor frecuencia. Puede tener una mayor variabilidad en los gradientes debido a que se calculan en función de un promedio más pequeño. Sin embargo, puede ser más estable y generalizable, ya que los pesos se actualizan con mayor frecuencia y los gradientes pueden reflejar mejor la información local.

La elección adecuada del número de épocas y el tamaño del batch depende del conjunto de datos, el modelo y los recursos disponibles. Por lo general, se necesita un proceso de ajuste y experimentación para encontrar el equilibrio correcto que permita un entrenamiento eficiente y una buena generalización del modelo.

## **2.4. Indicadores de error y métricas**

El estadístico utilizado para el cálculo de error en la etapa de entrenamiento es el MAE (Error Absoluto Medio).

### **2.4.1. MAE**

El error absoluto medio o MAE es una métrica que mide la magnitud de los errores en un conjunto de predicciones, sin considerar la dirección del error. Esto se utiliza para el entrenamiento de la red.

$$MAE = \frac{1}{n} \cdot \sum_{t=1}^n |y_t - \hat{y}_t|$$

Luego, los estadísticos utilizados para comparar las distintas redes son el MAPE (Error Porcentual Absoluto Medio) y el Accuracy (Precisión). Estos estadísticos se utilizan para medir los errores presentes en las predicciones.

### 2.4.2. MAPE

El error porcentual absoluto Medio o MAPE se utiliza para obtener un comparador universal entre las predicciones de la red y la demanda real. Calcula el promedio de los errores porcentuales absolutos entre las predicciones y los valores reales, lo que permite evaluar el error relativo de las predicciones independientemente de la escala de la variable de estudio.

$$MAPE = \frac{1}{n} \cdot \sum_{t=1}^n \frac{|y_t - \hat{y}_t| \cdot 100}{y_t}$$

### 2.4.3. Accuracy

Al igual que el MAPE se utiliza para obtener un comparador universal entre las predicciones de la red y la demanda real, siendo el valor complementario del MAPE.

$$Accuracy = 1 - MAPE$$

Dado que el MAPE representa el error relativo porcentual promedio entre las predicciones y los valores reales, restar este valor de 1 nos da la proporción de predicciones correctas.

## 2.5. Herramientas

Para programar la transformación de datos y las redes neuronales, se utilizó Colab, que es una plataforma en línea gratuita desarrollada por Google que permite a los usuarios escribir y ejecutar código en lenguaje Python a través del navegador web.

Python es un lenguaje de programación que permite utilizar diversas librerías desarrolladas por terceros, las cuales ofrecen una amplia gama de funcionalidades, dependiendo del enfoque o la tarea que se esté programando.

En el presente trabajo, se utilizan las siguientes librerías:

1. Pandas: Librería utilizada para manejar y transformar las bases de datos.
2. Numpy: Librería utilizada para manejar y transformar las bases de datos.
3. Matplotlib: Librería utilizada para graficar las series de tiempo
4. Seaborn: Librería utilizada para graficar las series de tiempo
5. Sklearn: Librería de algoritmos de aprendizaje automático.
6. TensorFlow: Librería utilizada para crear y entrenar modelos de aprendizaje automático.

### **2.5.1. TensorFlow**

TensorFlow es un ecosistema diseñado específicamente para facilitar y agilizar el trabajo con Machine Learning. TensorFlow ofrece diferentes niveles de abstracción, lo que permite adaptarse a las necesidades y conocimientos del usuario [16].

Una de las principales ventajas de TensorFlow es que utiliza Python para proporcionar una API conveniente, lo que permite construir aplicaciones de redes neuronales de manera más eficiente.

En TensorFlow, los desarrolladores tienen la capacidad de construir la estructura de las redes neuronales utilizando grafos, donde se representan las conexiones entre los distintos elementos utilizados.

Un grafo es una estructura matemática que representa las relaciones entre objetos, donde los objetos son denominados nodos y las relaciones entre ellos son representadas por enlaces.

Cada objeto del grafo representa una operación matemática específica, y las conexiones entre ellos se transforman en arreglos multidimensionales llamados tensores.

Los tensores son fundamentales en TensorFlow, ya que se utilizan para representar los datos en cada etapa del procesamiento. Estos arreglos multidimensionales pueden contener tanto datos de entrada como salidas intermedias, y permiten realizar cálculos eficientes y paralelos en el contexto del aprendizaje automático.

Por otro lado, una API (Application Programming Interface) es un conjunto de reglas y protocolos que permiten la comunicación entre diferentes aplicaciones. En el contexto de TensorFlow, se pueden utilizar diversas APIs para desarrollar redes neuronales de manera más sencilla. Un ejemplo es Keras, una API de aprendizaje profundo escrita en Python que se ejecuta en la capa superior de TensorFlow. Keras implementa técnicas de redes neuronales comúnmente utilizadas por los desarrolladores y se enfoca en facilitar la experimentación rápida con estas redes.

## **2.6. Elasticidad**

La Elasticidad es una medida que expresa cómo la cantidad demandada de un bien puede verse afectada por cambios en el precio de dicho bien. Se calcula mediante la fórmula de Elasticidad-Precio de la Demanda (EPD), que es el cambio porcentual en la cantidad demandada dividido por el cambio porcentual en el precio, esto se denota de la siguiente manera:

$$EPD = \frac{\frac{\Delta Q}{Q}}{\frac{\Delta P}{P}}$$

- $\Delta Q$  es el cambio en la cantidad demandada entre las semanas
- $Q$  es la cantidad demandada al inicio de la semana
- $\Delta P$  es el cambio en el precio entre las semanas
- $P$  es el precio al inicio de la semana

## 2.7. SAP IBP

SAP IBP es un sistema de planificación de la cadena de suministro que permite una integración y planificación eficiente de los procesos de negocio [17]. Esta plataforma toma las bases de datos proporcionadas por Agrosuper y realiza un conjunto de predicciones utilizando una gama de modelos de series de tiempo, seleccionando el mejor mediante la minimización de errores.

## 2.8. CRISP-DM

CRISP-DM es el acrónimo de "Cross-Industry Standard Process for Data Mining" (Proceso estándar multiindustrial para minería de datos). Se trata de un modelo de proceso de minería de datos que describe los enfoques comunes adoptados por los profesionales de la minería de datos. CRISP-DM proporciona una estructura que incluye seis fases principales:

1. Comprensión del negocio: Esta es la fase inicial que implica entender los objetivos y requisitos del proyecto desde una perspectiva empresarial.
2. Comprensión de los datos: En esta fase, se recopilan los datos y se examinan para familiarizarse con ellos, identificar la calidad de los datos, y descubrir subconjuntos iniciales para formar hipótesis.
3. Preparación de los datos: Los datos se limpian, transforman y enriquecen para el modelado. Esto puede incluir la solución de problemas de calidad de datos, la selección de variables y la transformación de variables.

4. Modelado: En esta fase, se seleccionan diversas técnicas de modelado y se calibran los parámetros a los datos preparados. En general, se escogen varios modelos y se evalúan para determinar cuál es el más efectivo para cumplir los objetivos del negocio.
5. Evaluación: Una vez construidos los modelos, se evalúan en el contexto de los objetivos empresariales. Esto implica determinar la utilidad y la validez del modelo en el contexto empresarial.
6. Implementación: Finalmente, el modelo se despliega en el entorno operacional, se monitoriza y se revisa para asegurar que está proporcionando los resultados esperados.

CRISP-DM es un enfoque iterativo, lo que significa que se necesite volver a fases anteriores varias veces antes de llegar a la implementación.

### 3. Metodología

La metodología utilizada para este proyecto es CRISP-DM [18] debido a su adaptabilidad y naturaleza cíclica. Su enfoque iterativo favorece la constante mejora de la red neuronal, a partir de la retroalimentación obtenida durante el desarrollo. Así, CRISP-DM proporciona un marco sólido para un desarrollo progresivo y ágil, con un objetivo enfocado en lograr predicciones cada vez más precisas en el campo del aprendizaje automático.

En un proyecto que implica el desarrollo de redes neuronales utilizando datos de una organización o entidad, en la práctica, se pueden identificar las siguientes etapas:

- **Comprensión del negocio:** Esta es la primera etapa, donde es crucial entender el contexto de la organización y las necesidades que se deben resolver. En esta fase, también se establecen los objetivos y se define un plan de proyecto detallado.
- **Comprensión de los datos:** Aquí se reúnen los datos iniciales y se procede a explorar y familiarizarse con ellos. Se lleva a cabo una evaluación de la calidad de los datos y se identifican las características relevantes que pueden ayudar a alcanzar los objetivos del proyecto.
- **Preparación de los datos:** En esta etapa se limpian y transforman los datos para su posterior análisis. Esto puede implicar la limpieza de datos incompletos o erróneos, la transformación de datos en un formato utilizable para el modelado y la creación de nuevas características a partir de las existentes.
- **Modelado:** Una vez que los datos están listos, se desarrolla un modelo de red neuronal. Este proceso implica seleccionar la arquitectura de red adecuada, entrenar el modelo con los datos preparados y ajustar los parámetros del modelo para mejorar su rendimiento.

- **Evaluación:** En esta fase, se evalúa el desempeño de la red. Se utilizan las métricas MAPE y Accuracy para medir la eficacia del modelo y su capacidad para predecir correctamente los resultados. Si los resultados son satisfactorios, se procede a la siguiente etapa. Si no lo son, se puede optar por revisar y ajustar el modelo o volver a las etapas anteriores para una mejor preparación de los datos.
- **Implementación:** Finalmente, una vez que el modelo ha sido evaluado y aprobado, se implementa en un entorno de producción. Aquí se realiza un seguimiento del rendimiento del modelo en tiempo real y se hacen ajustes según sea necesario.

En la siguiente ilustración se observa el diagrama del ciclo completo de la metodología:

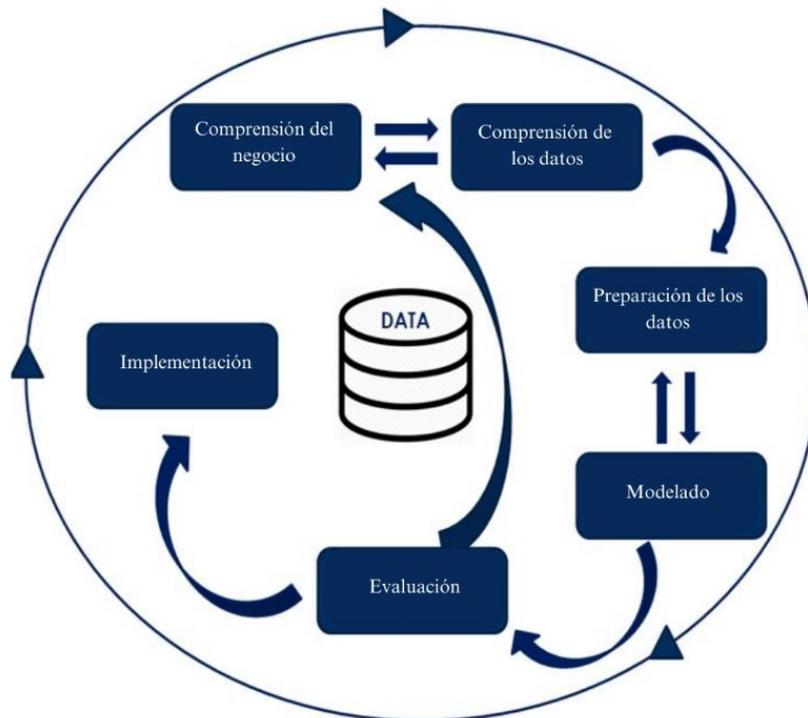


Ilustración 5: Diagrama de metodología CRISP DM

Fuente: Adaptación al español de Diagrama CRISP DM por Osvaldo G. Data Scientist en LinkedIn: Data Driven

Es importante destacar que, dada la metodología propuesta, completar un ciclo integral implica una considerable cantidad de trabajo, normalmente diseñado para equipos formados por ingenieros y científicos de datos. Por esta razón, para el trabajo de título, se busca ajustar esta metodología, excluyendo la fase de implementación de esta red neuronal, debido a su complejidad y el tiempo que se requiere para verificar y asegurar que esta red pueda reemplazar o competir con lo que se está utilizando actualmente para realizar las predicciones de demanda. En resumen, se adaptará la metodología CRISP-DM excluyendo la fase de implementación, ya que solo se busca tener evidencia de mejora en términos de precisión de predicción demanda.

## 4. Desarrollo del proyecto

El desarrollo del proyecto se inicia utilizando la metodología propuesta anteriormente, siguiendo una secuencia establecida que comienza con la comprensión y la preparación de los datos. Para esto se empieza con una limpieza de datos, luego la transformación de algunas variables temporales y la corrección de cualquier inconsistencia presente en los datos.

Después de la etapa de preparación de los datos, el proyecto prosigue con el modelado de la red neuronal. Se comienza con un caso base que implica una red LSTM (Long Short-Term Memory) que solo considera la demanda histórica.

Luego de haber establecido el caso base, se procede a modelar las otras redes LSTM utilizando diferentes sets de variables. Este proceso implicará la experimentación con diferentes combinaciones de variables de entrada y la configuración de la red neuronal.

Finalmente, se realiza la evaluación de los resultados obtenidos de las distintas redes con variables en comparación con el caso base inicialmente planteado. Esta etapa es crucial para determinar qué modelos y qué conjuntos de variables proporcionan las predicciones más precisas. La evaluación implica el uso de las métricas MAPE y Accuracy. Basándose en esta evaluación, se puede determinar cuál es el mejor modelo.

### 4.1. Comprensión y preparación de los datos

Las bases utilizadas para este proyecto fueron las siguientes:

#### 4.1.1 Base demanda y factura histórica:

Esta base de datos contiene 545,846 registros, que corresponden a datos semanales de dos sectores, desde mayo de 2018 hasta mayo de 2023. La base incluye 5 atributos de clasificación, 1 atributo descriptivo, 2 atributos temporales y 2 variables a considerar. A continuación, se describen estos atributos y variables:

1. Sector: Se refiere a la línea de negocio del producto específico, ya sea del sector cecinas o elaborados.
2. Tipo de cliente: Este atributo indica el canal de destino de venta, que puede ser Foodservice, Supermercados, Grandes Clientes o canal Tradicional.
3. Segmentación: Este atributo proporciona una segmentación de tipo de cliente a un nivel más detallado. Incluye supermercados específicos como Unimarc, Jumbo, Walmart, y también las distintas coberturas para canal Tradicional y Grandes clientes.

4. Tipo de producto: Indica el tipo de producto específico. Aquí se encuentran agrupaciones tales como salchichas, jamones, empanizados, entre otros.
5. Código SKU: Es el código unitario del producto específico.
6. Descripción: Es el nombre del producto específico.
7. ISOWeek: Enumeración de la semana ISO respectiva.

La ISOWeek es un sistema de numeración de semanas que sigue la norma ISO 8601. Esta norma establece que la primera semana del año es aquella que incluye al primer jueves del año, y las semanas comienzan en lunes.

8. Año: Es el año específico.
9. Demanda: Es la cantidad en kg de lo pedido del producto específico, dado los atributos anteriormente mencionados.
10. Factura: Es la cantidad en kg de lo vendido del producto específico, también determinado por los atributos anteriormente mencionados.

#### **4.1.2 Base de precio histórico:**

Esta base comparte casi todos los atributos con la base de demanda y factura histórica. La diferencia es que solo contiene 272,923 registros y en lugar de demanda y factura, tiene el precio unitario por producto específico como la variable clave.

La cantidad reducida de registros se debe a que a partir de mayo de 2020 se empezó a documentar un precio unitario por producto específico. Antes de eso, se calculaba un promedio, resultando en un precio promedio por cada SKU. Por este motivo, sólo se consideran los datos a partir de mayo de 2020.

#### **4.1.3 Base utilizada para el procesamiento:**

Las dos bases de datos se combinan mediante una llave única. En este proceso, aparecen algunos valores nulos para los precios de ciertos productos específicos.

Dado que el precio comienza en mayo de 2020, se considera este como el punto de inicio para la base de datos combinada. En consecuencia, la base combinada consta de 263,999 registros con un total de 5 atributos de clasificación, 1 atributo descriptivo, 2 atributos temporales y 3 variables.

Después de cargar la base de datos en Colab y de importar las bibliotecas pertinentes, se implementa un índice de fecha para cada registro en la base resultante. Para ello, se utilizan los valores de ISOWeek y Año, estableciendo como fecha el último día de cada ISOWeek para cada registro. Este proceso se realiza para facilitar el manejo de los datos.

#### 4.1.4 Base demanda real:

Esta base de datos se ocupará al final del proceso de los pronósticos, para realizar los cálculos de MAPE y Accuracy respectivamente.

### 4.2. Variables

La decisión de utilizar las variables factura y precio se fundamenta en que este proyecto es una fase piloto para evaluar si estas variables influyen en la predicción de la demanda. En este sentido, se han dejado de lado otras variables que podrían afectar la demanda, como los productos sustitutos y la dominancia en el mercado frente a la competencia, entre otros. La principal razón de esta omisión es que, para variables externas como la competencia, no se cuenta con registros exactos, y en el caso de los productos sustitutos, habría que hacer un levantamiento de una base de datos que obtenga estas clasificaciones por cada producto específico.

Sin embargo, contando con las variables de demanda y precio, es posible crear una variable que podría tener significancia, como la Elasticidad.

Dado que la Elasticidad se calcula por períodos, se puede considerar la semana -1 para observar el cambio. De lo contrario, la elasticidad sería constante. Es importante tener en cuenta que este enfoque permite capturar la variabilidad de la elasticidad a lo largo del tiempo.

Finalmente, se consideran la **factura**, el **precio** y la **elasticidad** como variables para las distintas variantes de redes LSTM a modelar.

#### 4.2.1 Características de las variables de estudio

Antes de utilizar los datos en el modelo de redes neuronales, se realizan algunos gráficos para examinar sus características a lo largo del tiempo e identificar posibles valores atípicos. Si se encuentran tales valores, podría ser necesario limitar y/o reemplazar algunos de los datos en la base. Este paso es esencial para asegurar la calidad y la relevancia de los datos que alimentarán el modelo.

## Cecinas:

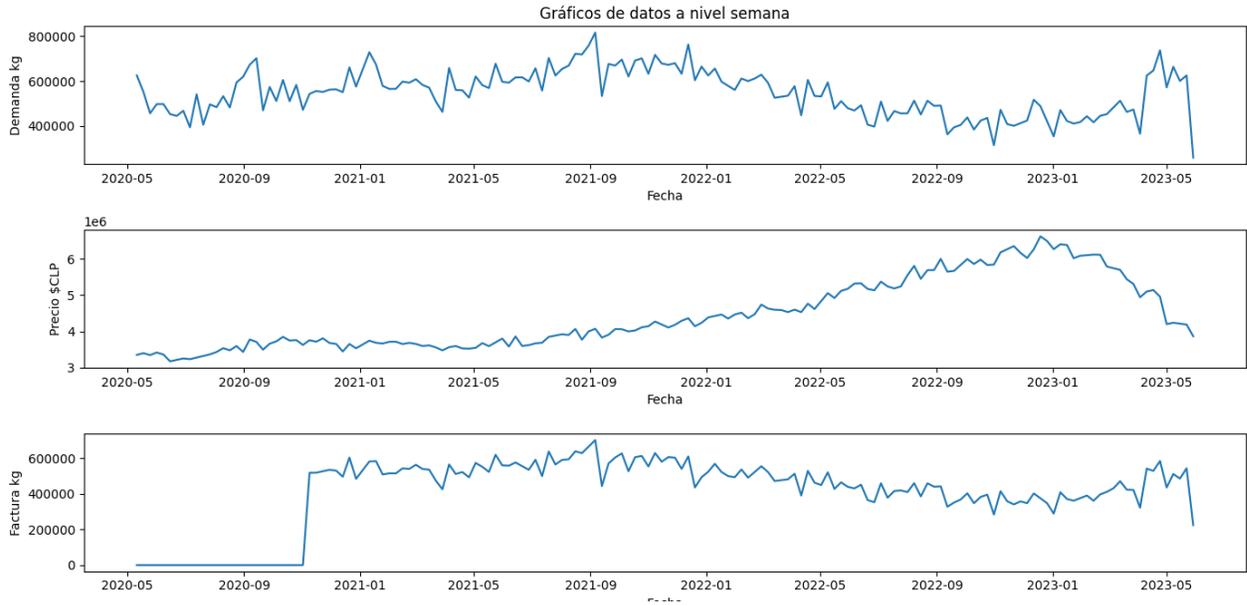


Ilustración 6: Gráficos de variables en el tiempo sector Cecinas

Fuente: Elaboración propia del estudiante utilizando Colab

## Elaborados:

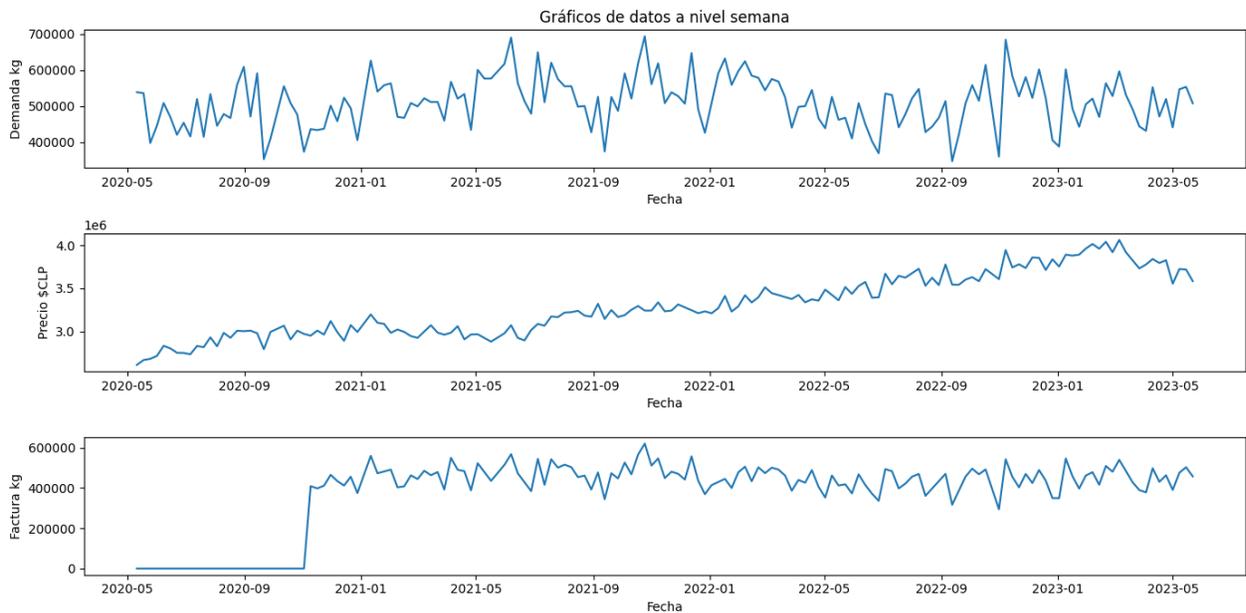


Ilustración 7: Gráficos de variables en el tiempo sector Elaborado

Fuente: Elaboración propia del estudiante utilizando Colab

A partir de estos gráficos, es evidente que existe una falta de datos en la variable Factura entre mayo y noviembre de 2020. Para tratar esta situación, se decidió completar los vacíos con el promedio de datos de Factura, variando dentro de un rango que se consideró razonable entre los 400 toneladas y 500 toneladas. El objetivo de esta estrategia es asegurar que las redes neuronales no experimenten problemas durante el aprendizaje a causa de esa falta de datos para esa variable.

También se puede observar que hay una correlación de la demanda con la factura, ya que la factura es lo que realmente se vendió (en kilogramos) apuntando a un pronóstico de demanda para esas fechas. Reiterando la aclaración de este concepto de factura, se tiene el valor de lo producido (en kilogramos) en cada fecha, la cual apunta a lo que se había pronosticado en términos de demanda dos semanas antes de esa fecha.

Asimismo, se puede observar que el precio ha ido incrementándose de manera constante a lo largo de los años, mientras que la demanda ha fluctuado dentro de un rango determinado,

### 4.3. Modelamiento

Para el modelamiento de una red neuronal LSTM se siguió este procedimiento:

- **Preparación de los datos:** Las redes LSTM requieren que los datos estén en un formato específico, usualmente en un arreglo tridimensional (muestras, pasos de tiempo, características). También se requirió normalizar los datos para mejorar la eficiencia de la red, donde se ocupó una función Min-Max de la librería Sklearn, que escala los datos al rango [0,1].
- **Definición del modelo LSTM:** Aquí es donde se define la arquitectura de la red. Se especifican las capas LSTM, las unidades dentro de esas capas y la función de activación. También se puede agregar una capa de salida, que tendrá tantas unidades como predicciones desee hacerse.
- **Compilación y entrenamiento del modelo:** En este paso, se define la función de pérdida que se utilizará (MAE) y el optimizador (Adam). Se alimenta los datos al modelo y se le permite aprender. Aquí también se definen el número de épocas y el tamaño del batch.
- **Evaluación del modelo de entrenamiento:** Una vez que el modelo ha sido entrenado, se puede evaluar su rendimiento en los datos de prueba. Este paso se omite por temas de tiempo, ya que habría que realizar un proceso de evaluación para cada entrenamiento de cada red LSTM que se hizo en este proyecto.
- **Predicción:** Finalmente, se usa el modelo para hacer predicciones sobre datos nuevos para después evaluar con métricas MAPE y Accuracy.

Durante la modelación de la red neuronal LSTM, surgieron complicaciones al presentar resultados a un nivel más detallado, es decir, Tipo de Cliente-Segmentación-Tipo de Producto-SKU. Esto se debió a la complejidad de la iteración y la automatización de la red, sin considerar que sería necesario emplear este mismo nivel de detalle a varias redes LSTM con sus distintos sets de variables. También hay que añadir que se requeriría una mayor cantidad de tiempo para procesar los datos y modelar las redes. Por ello, se optó por capturar la demanda a nivel detallado por semana, pero agrupando los datos al momento de utilizar la red neuronal. De esta manera, se pudo manipular los datos con mayor facilidad y hacer que el proceso de modelado fuera más rápido.

En base a este contexto, se llevó a cabo el modelado de una red LSTM de Caso Base que únicamente utiliza la demanda histórica para realizar su pronóstico.

Posteriormente, se desarrollaron diversas redes LSTM que consideran todas las combinaciones posibles de las variables y sus respectivas predicciones. Las distintas combinaciones se ilustran en el siguiente cuadro:

<b>Variables</b>	Red 1	Red 2	Red 3	Red 4	Red 5	Red 6	Red 7
Precio	Si	No	No	Si	Si	No	Si
Elasticidad	No	Si	No	Si	No	Si	Si
Factura	No	No	Si	No	Si	Si	Si

Tabla 1: Distribución de variables para cada red LSTM

También es importante mencionar que las predicciones se inician a partir del último dato disponible en la base utilizada para el modelamiento. Por lo tanto, la fecha inicial se establece en la primera semana de junio de 2023 sin considerar los días de mayo 2023, tomando el valor número 23 de la ISOWeek.

#### **4.4. Evaluación**

Teniendo ya los pronósticos, se utiliza una base que contiene la demanda real de las ISOWeek del mes de junio, para realizar los cálculos de MAPE y Accuracy respectivamente. Considerar que tener sobre un 80% de Accuracy es positivo, entre 60% y 70% es regular, y bajo del 60% ya es una mala precisión. Esto es según la modalidad que tiene en Agrosuper para supervisar sus métricas de Accuracy, dado esto se considerará estos valores para realizar los análisis respectivos de los resultados de los pronósticos de las redes neuronales.

Luego, los resultados alcanzados se expondrán en el próximo capítulo de este informe.

## 5. Resultados

Al inicio del proyecto de título, el estudiante tuvo la intención de utilizar el conglomerado de modelos empleados en el área de Planificación de la demanda. Sin embargo, estos no pueden ser modificados, ya que están implementados en una plataforma virtual llamada SAP IBP. Esto y en conjunto a los limitantes permisos que tiene un estudiante dentro la empresa, es por lo que no se puede acceder al código y su manipulan los datos dentro de esta plataforma.

Dada esta situación, se intentó utilizar de forma manual un conglomerado de modelos de series de tiempo. Sin embargo, el estudiante enfrentó dificultades al intentar modelar y capturar un valor similar a los de la demanda de productos procesados. Finalmente, se solicitó ayuda al área de Ciencias de Datos de Agrosuper. Este departamento ya había probado un piloto de uso de redes neuronales FeedForward, lo cual se tomó como inspiración para la creación de las redes LSTM que el estudiante utilizó para realizar las predicciones.

Dado este contexto, se muestran los resultados de los pronósticos de demanda proporcionados por todas las redes para ambos sectores de productos procesados, acompañados de su respectivo Accuracy.

### 5.1. Sector Elaborados

Se comienza con el caso base que solo considera la demanda histórica, este modelo base servirá para realizar las comparaciones con las otras redes. En la siguiente tabla se tiene los resultados de este Caso Base:

ISOWeek	Sector	Demanda real (kg)	Pronóstico (kg)	Accuracy
W23	Elaborados	639.629	510.178	79,7%
W24	Elaborados	596.851	526.690	88,2%
W25	Elaborados	504.721	533.256	94,3%
W26	Elaborados	496.423	535.604	92,1%

Tabla 2: Resultado red LSTM utilizando solo demanda histórica, sector elaborados

Como se puede ilustrar de esta tabla, no hay mucha variabilidad en términos de demanda pronosticada, pero si en términos de Accuracy, principalmente en la ISOWeek 23. De todas maneras, en casi todas las semanas se tiene un Accuracy positivo.

Luego se continúa con las variantes de las redes LSTM que consideran diferentes variables, para dar un ejemplo se tiene el siguiente gráfico del Accuracy de cada red para la ISOWeek 23, donde la primera columna es el caso base.

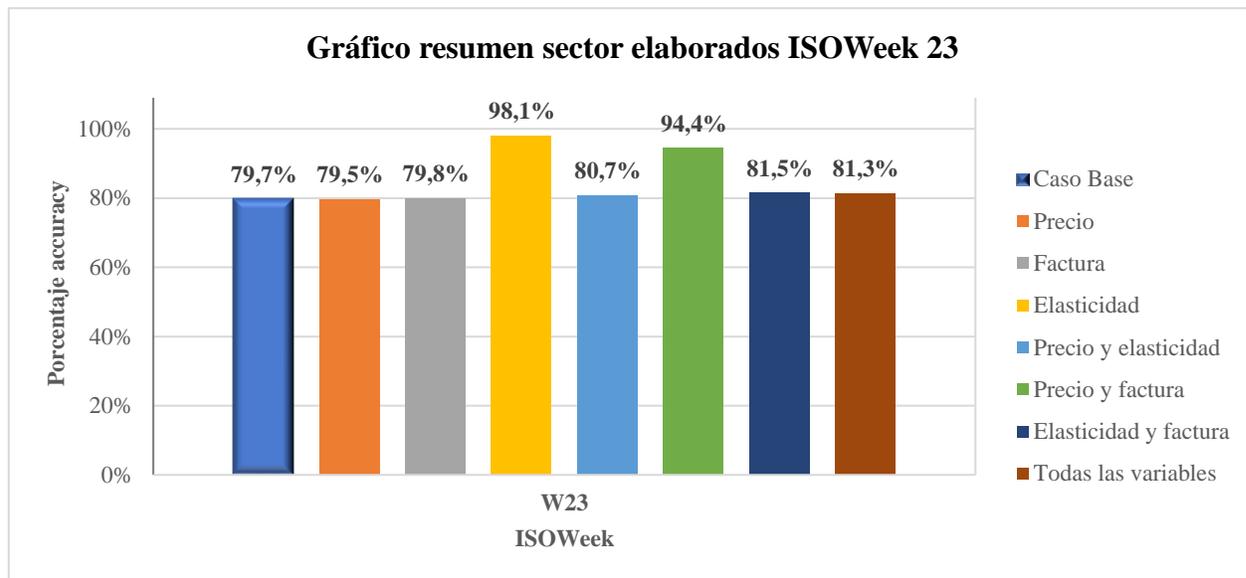


Ilustración 8: Gráfico resumen sector elaborados ISOWeek 23

Fuente: Elaboración propia del estudiante utilizando Excel

Se tiene en las Ilustraciones 14 a 16, los gráficos resumen de cada ISOWeek restante del mes de junio y también para observar las distintas combinaciones por separado se tienen las siguientes Ilustraciones 18 a 24 en el anexo. Los valores de los pronósticos y su demanda real están presentados en las Tablas 6 a 12 del anexo de este informe.

Con respecto a los resultados a nivel de ISOWeek se tiene lo siguiente:

ISOWeek 23: En casi todos los casos se presenta un Accuracy positivo por sobre el caso base, destacando los pronósticos de la red que solo considera elasticidad y la red que considera precio y factura (Ver Ilustración 8).

ISOWeek 24: En todos los casos se presenta un Accuracy por sobre el 80%, donde solo dos pronósticos destacaron por sobre el caso base, los cuales son de la red que solo considera elasticidad y la red que considera precio y factura.

ISOWeek 25: En casi todos los casos se presenta un Accuracy positivo por sobre el caso base, a excepción del pronóstico de la red que considera solo elasticidad y la red que considera precio y factura. En este punto, se puede observar inconsistencias en las predicciones de estas redes.

ISOWeek 26: Se obtienen resultados similares a la ISOWeek anterior, donde nuevamente los pronósticos de la red que considera solo elasticidad y la red que considera precio y factura tienen un Accuracy por debajo del caso base.

Luego, para determinar en primera instancia cuáles pronósticos apuntan a una mayor Accuracy, se realizó un filtro basado en el promedio semanal, del cual se obtiene el siguiente gráfico:

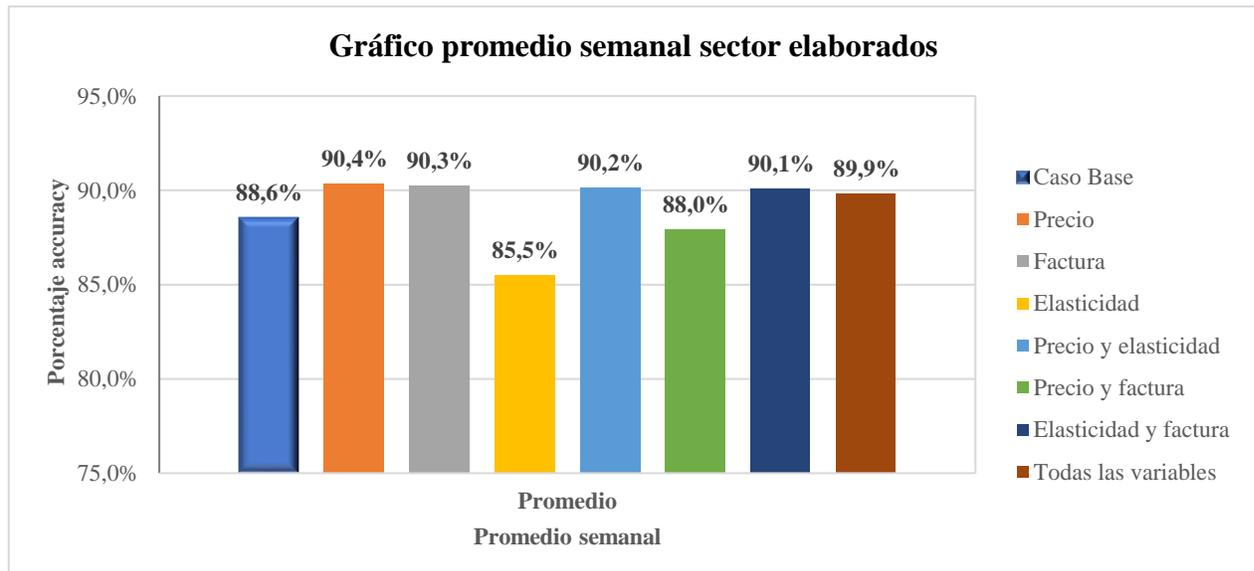


Ilustración 9: Gráfico promedio semanal sector elaborados

Fuente: Elaboración propia del estudiante utilizando Excel

Considerando este gráfico, se descartaría el pronóstico de la red que solo utiliza elasticidad y la red que utiliza precio y factura, debido a que ambas redes obtienen un Accuracy promedio menor al caso base.

Posteriormente, se aplicó un segundo filtro mediante una diferencia de Accuracy en relación con el caso base. Así, se decidió que los pronósticos que presenten una diferencia de Accuracy inferior al  $-5,5\%$  en alguna ISOWeek serán descartados. Dado esto se tiene la siguiente tabla:

ISOWeek	Caso base vs precio	Caso base vs factura	Caso base vs precio y elasticidad	Caso base vs elasticidad y factura	Caso base vs todas las variables
W23	-0,2%	0,1%	1,0%	1,8%	1,6%
W24	-3,0%	-2,6%	-1,6%	-0,8%	-1,0%
W25	4,9%	4,4%	3,2%	2,3%	2,5%
W26	5,4%	4,9%	3,7%	2,7%	2,0%

Tabla 3: Diferencia del Accuracy de las redes con variables vs Accuracy del caso base, sector elaborados

De esta tabla se tiene el siguiente gráfico, considerando también el primer filtro antes realizado.

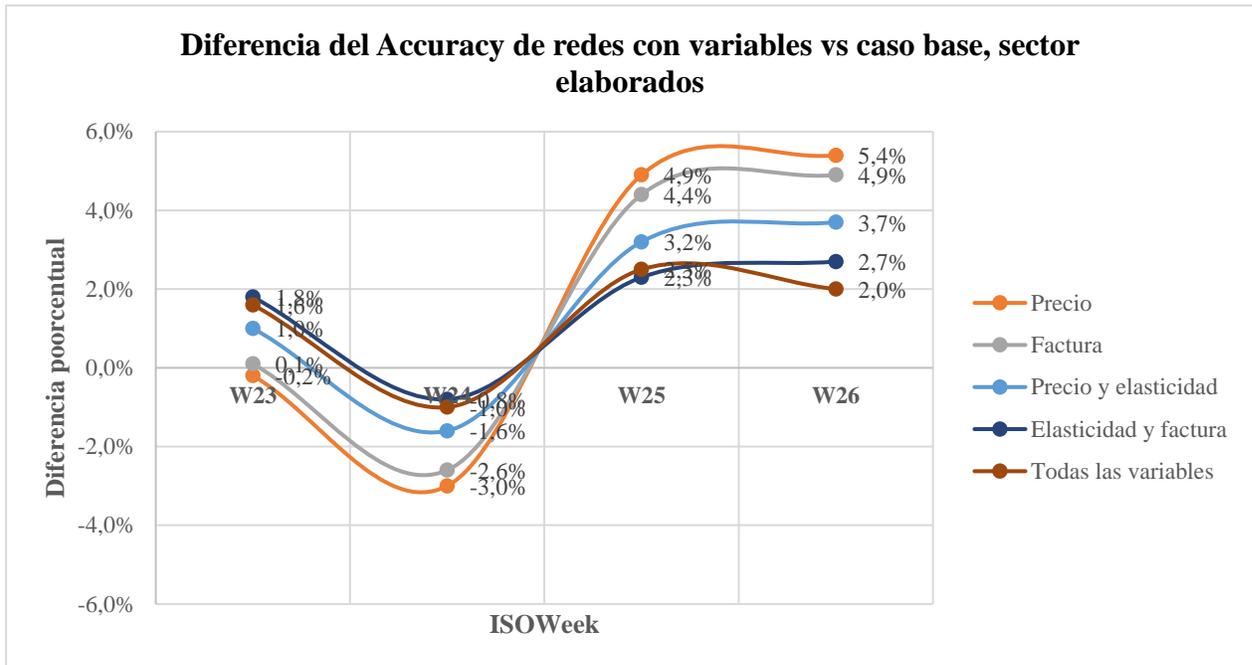


Ilustración 10: Diferencia del Accuracy de redes con variables vs caso base, sector elaborados

Fuente: Elaboración propia del estudiante utilizando Excel

Teniendo en cuenta los resultados presentados en la Tabla 3, se observa que los pronósticos de las redes seleccionadas muestran una diferencia de Accuracy superior al caso base. De todas maneras, existen casos en semanas particulares donde las diferencias están por debajo al caso base, pero no llegan a ser inferiores al -5,5%.

En base a los resultados de los pronósticos de demanda obtenidos de las diferentes redes neuronales LSTM, se puede apreciar que el pronóstico inicial del caso base ofrece un buen rendimiento, ya que en la mayoría de las semanas presenta una precisión superior al 80%. No obstante, al explorar diversas combinaciones de variables para la predicción de la demanda, se logra obtener un mayor grado de precisión, como se puede apreciar en la Ilustración 9, con la excepción de los pronósticos de la red que considera únicamente la elasticidad y la red que considera precio y factura, los cuales presentan inconsistencia destacable como se ilustra en la Tabla 3.

Tras aplicar el segundo filtro mencionado anteriormente, se constata que el pronóstico más preciso se logra con la red que solo utiliza el precio, ya que presenta un promedio de precisión del 90,4%. Este resultado se puede explicar por el crecimiento del precio, tal como se muestra en la Ilustración 7, donde la red es capaz de captar este incremento gracias a su capacidad de aprendizaje.

Cabe señalar que, a pesar de que los pronósticos de las demás redes son menos diferenciados que el caso base, estos ofrecen resultados positivos. De ello se deduce que, al incorporar nuevas variables, se logra una mayor precisión en comparación con el uso exclusivo de la demanda histórica.

## 5.2. Sector Cecinas

Análogamente, se comienza con el caso base que solo considera la demanda histórica. En la siguiente tabla se tiene los resultados de este caso base:

ISOWeek	Sector	Demanda real (kg)	Pronóstico (kg)	Accuracy
W23	Cecinas	583.984	462.035	79,1%
W24	Cecinas	626.789	474.730	75,7%
W25	Cecinas	566.459	455.437	80,4%
W26	Cecinas	530.987	461.917	87%

Tabla 4: Resultado red LSTM utilizando solo demanda histórica, sector cecinas

Como se puede ilustrar de esta tabla, no hay mucha variabilidad en términos de demanda pronosticada. Se puede señalar que en las primeras semanas los pronósticos son regulares en términos de Accuracy y en las últimas dos semanas se observan pronósticos positivos.

Luego se continúa con las variantes de las redes LSTM que consideran diferentes variables, para dar un ejemplo se tiene el siguiente gráfico del Accuracy de cada red para la ISOWeek 23, donde la primera columna es el caso base.

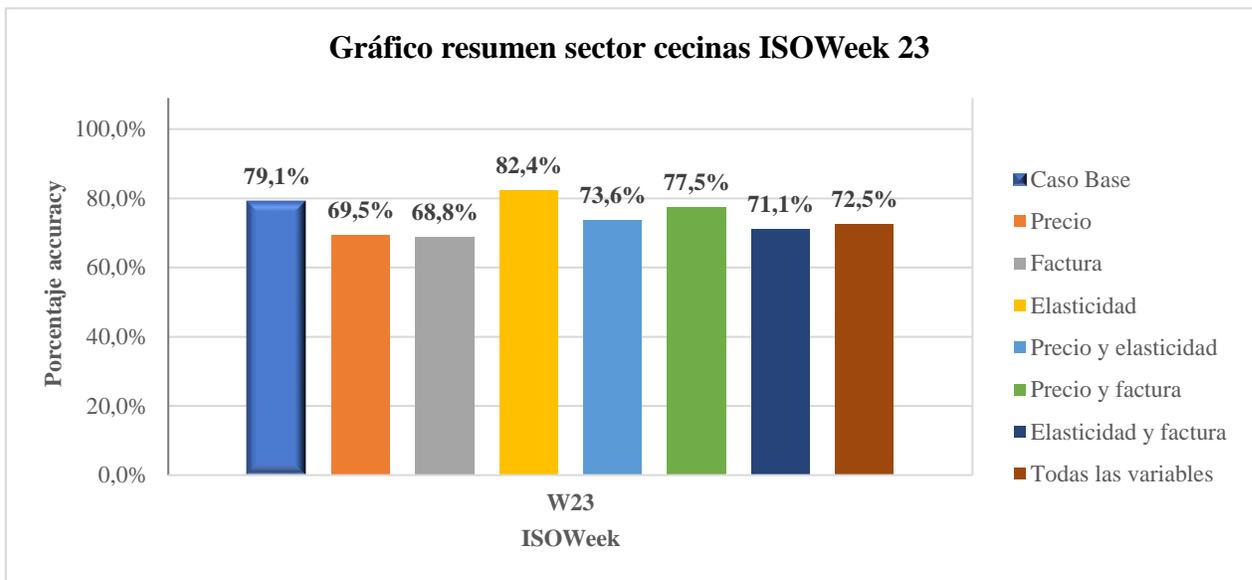


Ilustración 11: Gráfico resumen sector cecinas ISOWeek 23

Fuente: Elaboración propia del estudiante utilizando Excel

Se tiene en las Ilustraciones 24 a 26, los gráficos resumen de cada ISOWeek restante del mes de junio y también para observar las distintas combinaciones por separado se tienen las siguientes Ilustraciones 30 a 36 en el anexo. Los valores de los pronósticos y su demanda real están presentados en las Tablas 13 a 19 del anexo de este informe.

Con respecto a los resultados a nivel de ISOWeek se tiene lo siguiente:

**ISOWeek 23:** En casi todos los casos se presenta un Accuracy por debajo del caso base, solo un pronóstico está por sobre este caso base. Donde la red correspondiente a este pronóstico solo considera la variable de elasticidad (Ver Ilustración 11).

**ISOWeek 24:** En casi todos los casos se presenta un Accuracy positivo por sobre el caso base, donde solo un pronóstico está por debajo del caso base, el cual es el de la red que solo considera elasticidad (Ver Ilustración 26 del anexo de este informe).

**ISOWeek 25:** En casi todos los casos presentan un Accuracy positivo por sobre el caso base, a excepción de la red que considera solo elasticidad (Ver Ilustración 27 del anexo de informe).

**ISOWeek 26:** Se obtienen resultados similares a la ISOWeek anterior, donde nuevamente la red que considera solo elasticidad un Accuracy por debajo del caso base (Ver Ilustración 28 del anexo de informe).

Luego en primera instancia para acotar que pronósticos apuntan a un mayor Accuracy se realizó un promedio semanal, obteniendo el siguiente gráfico:

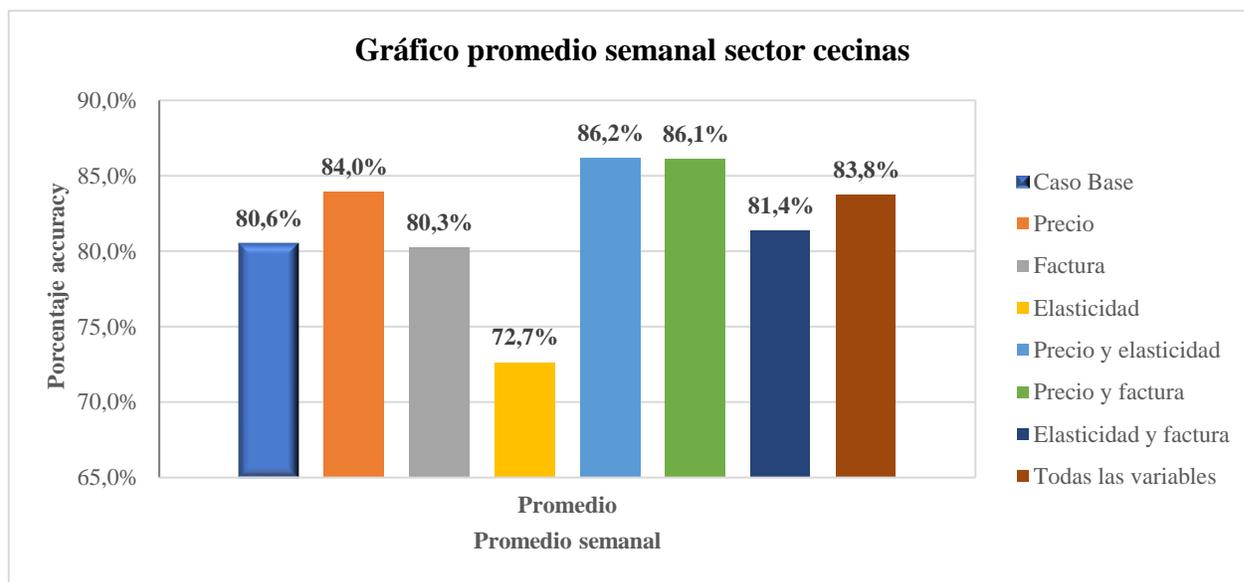


Ilustración 12: Gráfico promedio semanal sector cecinas

Fuente: Elaboración propia del estudiante utilizando Excel

Considerando este gráfico, se descartaría el pronóstico de la red que solo utiliza elasticidad y la red que solo utiliza factura, debido a que ambas redes obtienen un Accuracy promedio menor al caso base.

Posteriormente, se aplicó un segundo filtro mediante una diferencia porcentual en relación con el caso base. Así, se decidió que los pronósticos que presenten una diferencia porcentual inferior al  $-5,5\%$  en alguna ISOWeek serán descartados. Dado esto se tiene la siguiente tabla:

ISOWeek	Caso base vs precio	Caso base vs precio y elasticidad	Caso base vs precio y factura	Caso base vs elasticidad y factura	Caso base vs todas las variables
W23	-9,6%	-5,5%	-1,6%	-8,0%	-6,6%
W24	1,2%	5,1%	4,0%	-0,9%	1,3%
W25	11,9%	15,4%	9,9%	5,8%	8,6%
W26	10,2%	7,5%	9,9%	6,3%	9,5%

Tabla 5: Diferencia del Accuracy de las redes con variables vs Accuracy del caso base, sector cecinas

De esta tabla se tiene el siguiente gráfico, considerando también el primer filtro antes realizado.

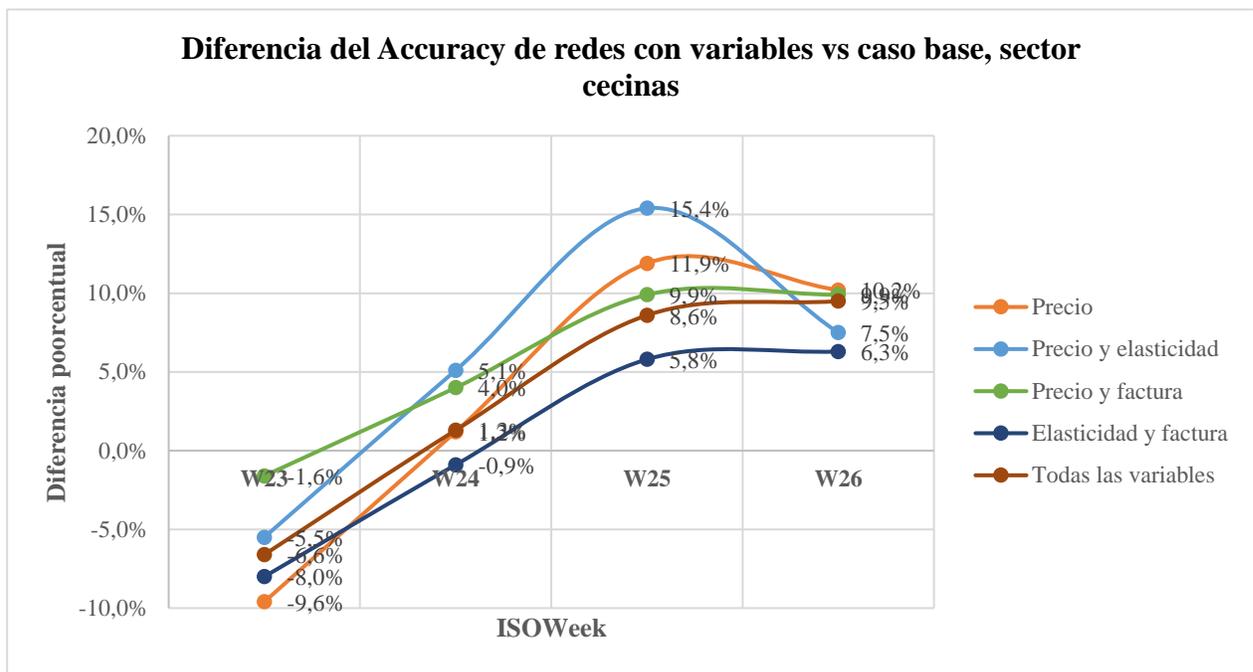


Ilustración 13: Diferencia del Accuracy de redes con variables vs caso base, sector cecinas

Fuente: Elaboración propia del estudiante utilizando Excel

Teniendo en cuenta los resultados presentados en la Tabla 5, se observa que los pronósticos de las redes seleccionadas muestran una diferencia de Accuracy superior al caso base a partir de la ISOWeek 24. De todas maneras, existen casos en la ISOWeek 23 donde las diferencias están por debajo al caso base llegando a ser inferiores al -5,5%, por lo que se descartaría la mayoría de los pronósticos, exceptuando el pronóstico de la red que considera precio y elasticidad, y el pronóstico de la red que considera precio y factura.

En base a los resultados de los pronósticos de demanda obtenidos de las diferentes redes neuronales LSTM, se puede apreciar que el pronóstico inicial del caso base ofrece un rendimiento regular, ya que presenta una precisión apenas del 80% en promedio. No obstante, al explorar diversas combinaciones de variables para la predicción de la demanda, se logra obtener un mayor grado de precisión, como se puede apreciar en la Ilustración 12, con la excepción del pronóstico de la red que considera únicamente la elasticidad. El cual presenta una inconsistencia destacable como se ilustra en la Tabla 5.

Después de aplicar el segundo filtro mencionado previamente, se evidencia que el pronóstico más preciso se tiene con la red que utiliza las variables de precio y la elasticidad, la cual presenta una precisión promedio del 86,2%. De manera análoga al sector anterior, este resultado se puede justificar por el incremento del precio, tal como se indica en la Ilustración 6. Sin embargo, la diferencia reside en que la demanda presenta una elasticidad más alta en comparación con el sector de productos elaborados, ya que, al aumentar el precio, la demanda disminuye de forma

más marcada. En este sentido, la red es capaz de captar esta compensación entre precio y elasticidad.

Con la obtención de estos resultados en ambos sectores, se logra satisfacer tanto el objetivo general como los objetivos específicos establecidos. Esto se debe a que se alcanzó un nivel de pronóstico coherente, que se asemeja a los valores de la demanda real para esas semanas, mediante la utilización de una red neuronal como herramienta de predicción. Además, la inclusión de variables ha demostrado conllevar una mejora en la precisión de la estimación de la demanda para ambos sectores en cuestión.

## 6. Conclusiones

### 6.1. Conclusión y recomendaciones

Es importante señalar que los pronósticos de las redes que incluyen variables indican una mayor precisión respecto al caso base, ya que estos proporcionan resultados positivos. De esto se infiere que, al incorporar nuevas variables, se obtiene una mayor precisión en comparación con el uso exclusivo de la demanda histórica, esto debido a que incluir múltiples variables en un modelo de predicción de demanda generalmente mejora la precisión de los pronósticos por varias razones, tales como captura de la complejidad del mercado, considerando la estacionalidad y las tendencias del sector de productos procesados. De lo anterior, el autor N. Gregory Mankiw en su libro “Principles of Economics” [19], específicamente en el capítulo 2 “How Markets Work” señala al comportamiento de la demanda a estas variables, entre los cuales también podrían influir las variables tales como factores económicos, competencia, comportamiento del consumidor e interacciones entre estas variables.

Esto sugiere que es recomendable considerar y utilizar variables adicionales al momento de modelar y pronosticar la demanda. La incorporación de variables como el precio, la factura, la elasticidad u otras relevantes puede proporcionar información adicional y capturar mejor las complejidades y las interacciones presentes en los datos de demanda.

Es importante destacar que cada caso y conjunto de variables puede tener sus particularidades y requerir ajustes específicos en los parámetros y en el proceso de modelado. Sin embargo, en general, la exploración y utilización de nuevas variables puede llevar a una mejora en la precisión de los pronósticos de demanda.

Luego, se pueden identificar varios factores que podrían contribuir a las inconsistencias observadas:

1. Ajuste de parámetros: Es posible que sea necesario ajustar los parámetros del modelo LSTM, como el número de capas LSTM, las unidades dentro de esas capas, la función de activación, la función de pérdida, el optimizador, el número de épocas y el tamaño del

batch. Realizar pruebas y ajustes adicionales en estos parámetros podrían mejorar el rendimiento y la precisión del modelo.

2. Utilización de otras variables: Es importante considerar que existen otras variables que podrían tener un efecto significativo en la demanda y que no se han incluido en el análisis. Por ejemplo, la competencia en el mercado o la presencia de productos sustitutos podrían tener un impacto relevante en los resultados de pronóstico. Sin embargo, debido a limitaciones de disponibilidad de datos y alcance del proyecto, estas variables no se han considerado en este estudio.
3. Tiempo de ensayo y error: Es importante tener en cuenta que el proceso de desarrollo y ajuste de modelos de pronóstico de demanda implica un tiempo de ensayo y error. Los resultados obtenidos hasta el momento pueden ser parte de un proceso iterativo en el cual se van realizando ajustes y mejoras en el modelo a medida que se obtienen nuevos resultados y se adquiere más experiencia en el análisis de la demanda. Dado esto, se recomienda realizar un nivel de agregación de manera paulatina, es decir, de lo más general (sector) a lo más particular (nivel SKU), para que en cada nivel se tenga predicciones coherentes en términos de demanda.
4. Utilización de otros modelos de series de tiempo: Aunque se haya utilizado el modelo LSTM en este proyecto, existen otros modelos de series de tiempo disponibles para el pronóstico de la demanda. Sin embargo, en este caso en particular, no se ha podido acceder a otros modelos debido a limitaciones en los permisos y al uso de una plataforma específica (SAP IBP) que utiliza su propia metodología de pronóstico.
5. Nivel de agregación de los datos: Durante el proceso de modelado de la red LSTM, se encontraron complicaciones al intentar procesar y presentar resultados a un nivel más detallado, considerando segmentaciones específicas. Esto se debe a la complejidad y tiempo requeridos para iterar y automatizar la red en cada nivel de detalle. Por lo tanto, se optó por capturar la demanda a nivel agregado por semana, lo que permitió un procesamiento más eficiente de los datos y aceleró el proceso de modelado, pero a consecuencia de no tener un nivel de estabilidad deseado de pronósticos.

## 7. Aprendizajes

Durante el proceso de investigación y desarrollo, se adquirió valiosos conocimientos sobre el uso de redes neuronales y su aplicación en la integración de múltiples variables en modelos de pronóstico. Estos aprendizajes no solo resultan fundamentales para la culminación de la carrera profesional, sino que también prometen ser una herramienta valiosa en futuros trabajos como ingeniero.

La experiencia adquirida en este proyecto de investigación ofrece oportunidades significativas para la aplicación práctica de las redes neuronales en diversos ámbitos. En particular, se

considera que estos modelos podrían ser beneficiosos en proyectos relacionados con la planificación y gestión de la cadena de suministro, la toma de decisiones estratégicas y la optimización de procesos logísticos. Estos campos de aplicación representan desafíos importantes en el mundo empresarial actual, y las técnicas avanzadas de pronóstico, como las redes neuronales, tienen el potencial de mejorar significativamente la eficiencia y la rentabilidad de las operaciones.

Además de los logros obtenidos, quedan algunas cuestiones pendientes que podrían ser objeto de futuras investigaciones. Por ejemplo, se identifica la necesidad de analizar en profundidad el impacto de la sobreestimación y subestimación de la demanda en la logística de productos procesados, y cómo estas discrepancias pueden afectar negativamente a los negocios. También el efecto que tiene la competencia, ya que un ejemplo es lo que se señaló en los antecedentes generales de este informe sobre el sector de productos de cecinas, el cual tiene un 7% de participación de mercado en Chile y cómo afectaría esto en las predicciones y su planificación productiva. En este sentido, esta área de estudio ofrece un campo fértil para investigaciones posteriores y podría proporcionar información valiosa para la toma de decisiones en la planificación de la cadena de suministro.

Culminando con los aprendizajes, se reconoce la importancia de mejorar sus habilidades de gestión del tiempo y la planificación de proyectos. A través de la experiencia adquirida en este proyecto, se ha vuelto evidente que la optimización del tiempo y la organización son factores cruciales para la consecución exitosa de objetivos académicos y profesionales.

## 8. Bibliografía

- [1] Banco Central de Chile. (2022). *Producto Interno Bruto Regional Año 2022*. Recuperado el 30 de Junio de 2023, de <https://www.bcentral.cl/contenido/-/detalle/producto-interno-bruto-regional-2022>
- [2] Asociación de Exportadores de Carnes de Chile A.G., ChileCarne. (2022). *Chile Carne, Exportadores de Carne de Chile*. Recuperado el 23 de Mayo de 2023, de Chile Carne, La industria en cifras: <https://www.chilecarne.cl/la-industria-en-cifras/>
- [3] Oficina de Estudios y Políticas Agrarias ODEPA. (2022). *ODEPA*. Recuperado el 23 de Mayo de 2023, de ODEPA Rubros | Carnes: <https://www.odepa.gob.cl/rubros/carnes>
- [4] AGROSUPER S.A. (s.f.). *AGROSUPER S.A.* Recuperado el 23 de Mayo de 2023, de AGROSUPER, Gobierno corporativo, nuestra empresa: <https://www.agrosuper.cl/gobierno-corporativo/nuestra-empresa/>
- [5] The Canvas Group. (05 de Agosto de 2020). *The Canvas Group*. Recuperado el 23 de Mayo de 2023, de Por qué no basta con hablar de la misión y la visión de tu empresa | Agrosuper: <https://www.canvasgroup.com/no-basta-con-hablar-de-mision-y-vision/>
- [6] AGROSUPER S.A. (2022). *Reporte Integrado AGROSUPER 2022*. Recuperado el 23 de Mayo de 2023, de <https://www.agrosuper.cl/wp-content/uploads/2023/04/Reporte-Agrosuper-2022.pdf>
- [7] LIS DATA SOLUTIONS. (s.f.). *Predicción de la demanda para optimizar la planificación de la cadena de suministro*. Recuperado el 30 de Junio de 2023, de <https://www.lisdatasolutions.com/es/blog/prediccion-de-la-demanda-para-optimizar-la-planificacion-de-la-cadena-de-suministro/>
- [8] Ch., C. M. (2013). *Comparación de los métodos de series de tiempo y redes neuronales*. Universidad Nacional Agraria La Molina, Lima - Perú. doi:<http://dx.doi.org/10.21704/ac.v75i2.960>
- [9] Khan Academy. (2015). *Khan Academy*. Recuperado el 27 de Junio de 2023, de ¿Qué factores modifican la demanda?: <https://es.khanacademy.org/economics-finance-domain/microeconomics/supply-demand-equilibrium/demand-curve-tutorial/a/what-factors-change-demand>
- [10] Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (1970). *TIME SERIES ANALYSIS Forecasting and Control* (Quinta ed.). (D. J. Balding, N. A. Cressie, G. M. Fitzmaurice, G. H. Givens, H. Goldstein, G. Molenberghs, . . . S. Weisberg, Edits.) WILEY SERIES IN PROBABILITY AND STATISTICS. Recuperado el 31 de Mayo de 2023, de [http://repo.darmajaya.ac.id/4781/1/Time%20Series%20Analysis\\_%20Forecasting%20and%20Control%20%28%20PDFDrive%20%29.pdf](http://repo.darmajaya.ac.id/4781/1/Time%20Series%20Analysis_%20Forecasting%20and%20Control%20%28%20PDFDrive%20%29.pdf)
- [11] Haoyi Zhou, S. Z. (2020). *Informer: Beyond Efficient Transformer for Long Sequence*. Beijing Guowang Fuda Science & Technology Development Company. Recuperado el 15 de Junio de 2023, de [https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwj3q5X0geb\\_AhXqIrkGHVwKAYIQFnoECBIQAQ&url=https%3A%2F%2Ffarxiv.org%2Fabs%2F2012.07436&usq=AOvVaw0OYq8aXLqCAvWRMRmqLY3G&opi=89978449](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwj3q5X0geb_AhXqIrkGHVwKAYIQFnoECBIQAQ&url=https%3A%2F%2Ffarxiv.org%2Fabs%2F2012.07436&usq=AOvVaw0OYq8aXLqCAvWRMRmqLY3G&opi=89978449)
- [12] AWS Amazon. (2023). *¿Qué es una red neuronal?* Recuperado el 30 de Junio de 2023, de <https://aws.amazon.com/es/what-is/neural-network/>

- [13] DeepAI. (s.f.). *DeepAI*. Recuperado el 15 de Junio de 2023, de What is a Feed Forward Neural Network?: <https://deepai.org/machine-learning-glossary-and-terms/feed-forward-neural-network>
- [14] Hasim Sak, A. S. (2014). *Long Short-Term Memory Recurrent Neural Network Architectures*. Recuperado el 15 de Junio de 2023, de <https://static.googleusercontent.com/media/research.google.com/es//pubs/archive/43905.pdf>
- [15] Diederik P. Kingma, J. B. (2015). *ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION*. Recuperado el 30 de Junio de 2023, de <https://arxiv.org/abs/1412.6980>
- [16] TensorFlow. (2023). *Introducción a TensorFlow*. Recuperado el 30 de Junio de 2023, de <https://www.tensorflow.org/learn?hl=es-419>
- [17] Kepczynski, R., Dimofte, A., Jandhyala, R., Sankaran, G., & Boyle, A. (2019). *Implementing Integrated Business Planning: A Guide Exemplified With Process Context and SAP IBP Use Cases*. (S. I. AG, Ed.) Springer International Publishing AG. Recuperado el 23 de Mayo de 2023
- [18] Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0 Step-by-step data mining guide*. NCR Systems Engineering Copenhagen; DaimlerChrysler AG; SPSS Inc.; OHRA Verzekeringen en Bank Groep B.V. Recuperado el 31 de Mayo de 2023, de <http://www.statoo.com/CRISP-DM.pdf>
- [19] N. Gregory Mankiw (2015), *Principles of Economics*, Cengage Learning , Octava Edición, Recuperado el 30 de Agosto de 2023, de [https://library.samdu.uz/files/308daa6da40e2853934a968f3e30d96c\\_Principles%20of%20Economics.pdf](https://library.samdu.uz/files/308daa6da40e2853934a968f3e30d96c_Principles%20of%20Economics.pdf)

# ANEXOS

## Anexo A: Tablas

### Sector Elaborados

#### Resultados LSTM variable precio, sector elaborados

ISOWeek	Sector	Demanda real	Pronóstico	Accuracy
W23	Elaborados	639.629 (kg)	508.582 (kg)	79,5%
W24	Elaborados	596.851	508.645	85,2%
W25	Elaborados	504.721	508.664	99,2%
W26	Elaborados	496.423	508.669	97,5%

Tabla 6: Resultado red LSTM utilizando variable precio, sector elaborados

#### Resultados LSTM variable factura, sector elaborados

ISOWeek	Sector	Demanda real	Pronóstico	Accuracy
W23	Elaborados	639.629 (kg)	510.980 (kg)	79,8%
W24	Elaborados	596.851	511.088	85,6%
W25	Elaborados	504.721	511.110	98,7%
W26	Elaborados	496.423	511.114	97%

Tabla 7: Resultado red LSTM utilizando variable factura, sector elaborados

## Resultados LSTM variable elasticidad, sector elaborados

ISOWeek	Sector	Demanda real	Pronóstico	Accuracy
W23	Elaborados	639.629 (kg)	627.971 (kg)	98,1%
W24	Elaborados	596.851	627.978	94,7%
W25	Elaborados	504.721	627.979	75,7%
W26	Elaborados	496.423	627.979	73,5%

Tabla 8: Resultado red LSTM utilizando variable elasticidad, sector elaborados

## Resultados LSTM variables precio y elasticidad, sector elaborados

ISOWeek	Sector	Demanda real	Pronóstico	Accuracy
W23	Elaborados	639.629 (kg)	516.453 (kg)	80,7%
W24	Elaborados	596.851	516.904	86,6%
W25	Elaborados	504.721	517.019	97,5%
W26	Elaborado	496.423	517.048	95,8%

Tabla 9: Resultado red LSTM utilizando variables precio y elasticidad, sector elaborados

## Resultados LSTM variables precio y factura, sector elaborados

ISOWeek	Sector	Demanda real	Pronóstico	Accuracy
W23	Elaborados	639.629 (kg)	603.837 (kg)	94,4%
W24	Elaborados	596.851	603.936	98,8%
W25	Elaborados	504.721	603.939	80,3%
W26	Elaborados	496.423	603.939	78,3%

Tabla 10: Resultado red LSTM utilizando variables precio y factura, sector elaborados

## Resultados LSTM variables elasticidad y factura, sector elaborados

ISOWeek	Sector	Demanda real	Pronóstico	Accuracy
W23	Elaborados	639.629 (kg)	521.402 (kg)	81,5%
W24	Elaborados	596.851	521.784	87,4%
W25	Elaborados	504.721	521.853	96,6%
W26	Elaborados	496.423	521.866	94,8%

Tabla 11: Resultado red LSTM utilizando variables elasticidad y factura, sector elaborados

## Resultados LSTM variables precio, factura y elasticidad, sector elaborados

ISOWeek	Sector	Demanda real	Pronóstico	Accuracy
W23	Elaborados	639.629 (kg)	520.174 (kg)	81,3%
W24	Elaborados	596.851	520.556	87,2%
W25	Elaborados	504.721	520.629	96,8%
W26	Elaborados	496.423	520.642	94,1%

Tabla 12: Resultado red LSTM utilizando variables precio, factura y elasticidad, sector elaborados

## Sector Cecinas

### Resultados LSTM variable precio, sector cecinas

ISOWeek	Sector	Demanda real (kg)	Pronóstico (kg)	Accuracy
W23	Cecinas	583.984	405.818	69,5%
W24	Cecinas	626.789	481.956	76,9%
W25	Cecinas	566.459	523.066	92,3%
W26	Cecinas	530.987	545.657	97,2%

Tabla 13: Resultado red LSTM utilizando variable precio, sector cecinas

### Resultados LSTM variable factura, sector cecinas

ISOWeek	Sector	Demanda real (kg)	Pronóstico (kg)	Accuracy
W23	Cecinas	583.984	402.007	68,8%
W24	Cecinas	626.789	460.288	73,4%
W25	Cecinas	566.459	484.904	85,6%
W26	Cecinas	530.987	495.428	93,3%

Tabla 14: Resultado red LSTM utilizando variable factura, sector cecinas

### Resultados LSTM variable elasticidad, sector cecinas

ISOWeek	Sector	Demanda real (kg)	Pronóstico (kg)	Accuracy
W23	Cecinas	583.984	686.548	82,4%
W24	Cecinas	626.789	742.136	81,6%
W25	Cecinas	566.459	748.970	67,8%
W26	Cecinas	530.987	749.803	58,8%

Tabla 15: Resultado red LSTM utilizando variable elasticidad, sector cecinas

### Resultados LSTM variables precio y elasticidad, sector cecinas

ISOWeek	Sector	Demanda real (kg)	Pronóstico (kg)	Accuracy
W23	Cecinas	583.984	429.695	73,6%
W24	Cecinas	626.789	506.643	80,8%
W25	Cecinas	566.459	542.934	95,8%
W26	Cecinas	530.987	560.350	94,5%

Tabla 16: Resultado red LSTM utilizando variables precio y elasticidad, sector cecinas

### Resultados LSTM variables precio y factura, sector cecinas

ISOWeek	Sector	Demanda real (kg)	Pronóstico (kg)	Accuracy
W23	Cecinas	583.984	452.543	77,5%
W24	Cecinas	626.789	499.793	79,7%
W25	Cecinas	566.459	511.605	90,3%
W26	Cecinas	530.987	514.571	96,9%

Tabla 17: Resultado red LSTM utilizando variables precio y factura, sector cecinas

### Resultados LSTM variables elasticidad y factura, sector cecinas

ISOWeek	Sector	Demanda real (kg)	Pronóstico (kg)	Accuracy
W23	Cecinas	583.984	415.169	71,1%
W24	Cecinas	626.789	469.050	74,8%
W25	Cecinas	566.459	488.272	86,2%
W26	Cecinas	530.987	495.197	93,3%

Tabla 18: Resultado red LSTM utilizando variables elasticidad y factura, sector cecinas

### Resultados LSTM variables precio, factura y elasticidad, sector cecinas

ISOWeek	Sector	Demanda real (kg)	Pronóstico (kg)	Accuracy
W23	Cecinas	583.984	423.104	72,5%
W24	Cecinas	626.789	482.417	77%
W25	Cecinas	566.459	504.377	89%
W26	Cecinas	530.987	512.563	96,5%

Tabla 19: Resultado red LSTM utilizando variables precio, factura y elasticidad, sector cecinas

## Anexo B: Gráficos

### Sector Elaborados

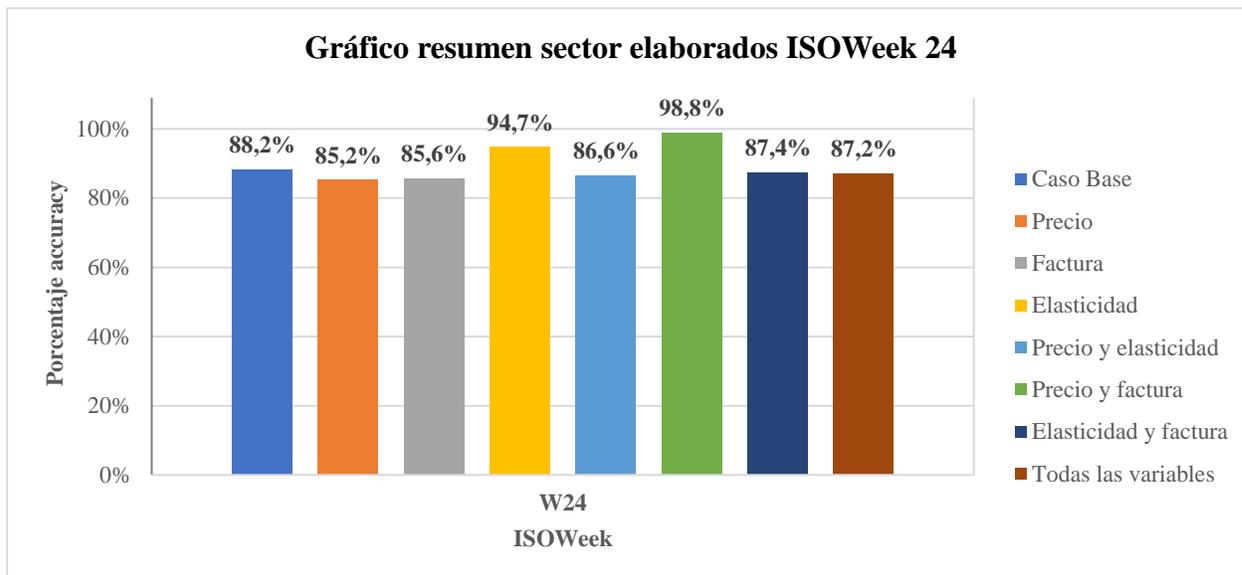


Ilustración 14: Gráfico resumen sector elaborados ISOWeek 24

Fuente: Elaboración propia del estudiante utilizando Excel

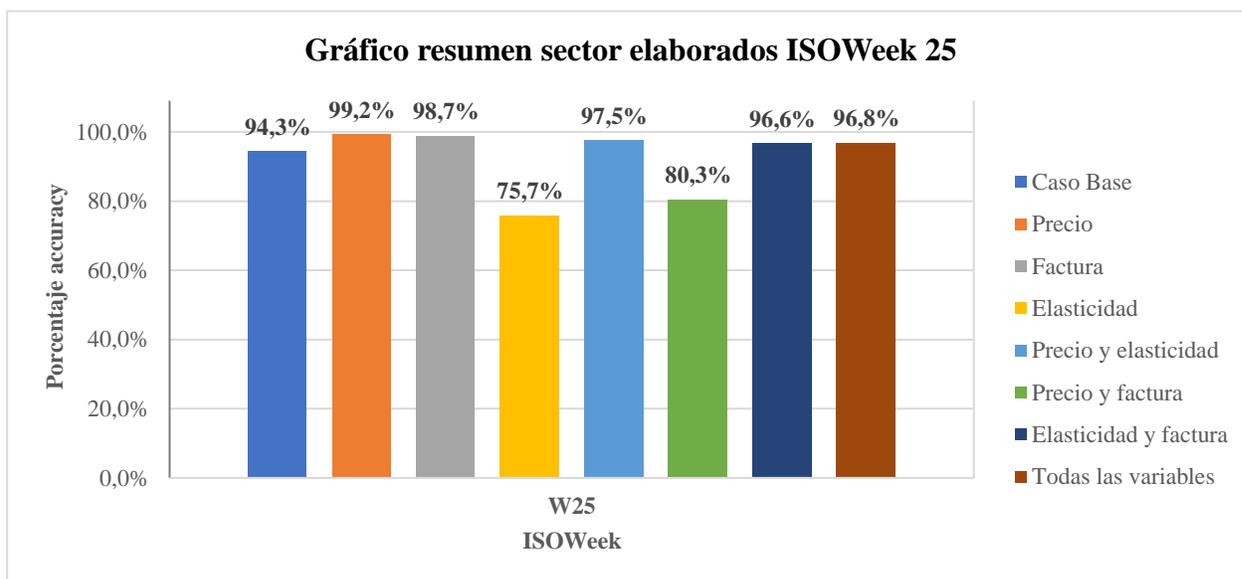


Ilustración 15: Gráfico resumen sector elaborados ISOWeek 25

Fuente: Elaboración propia del estudiante utilizando Excel

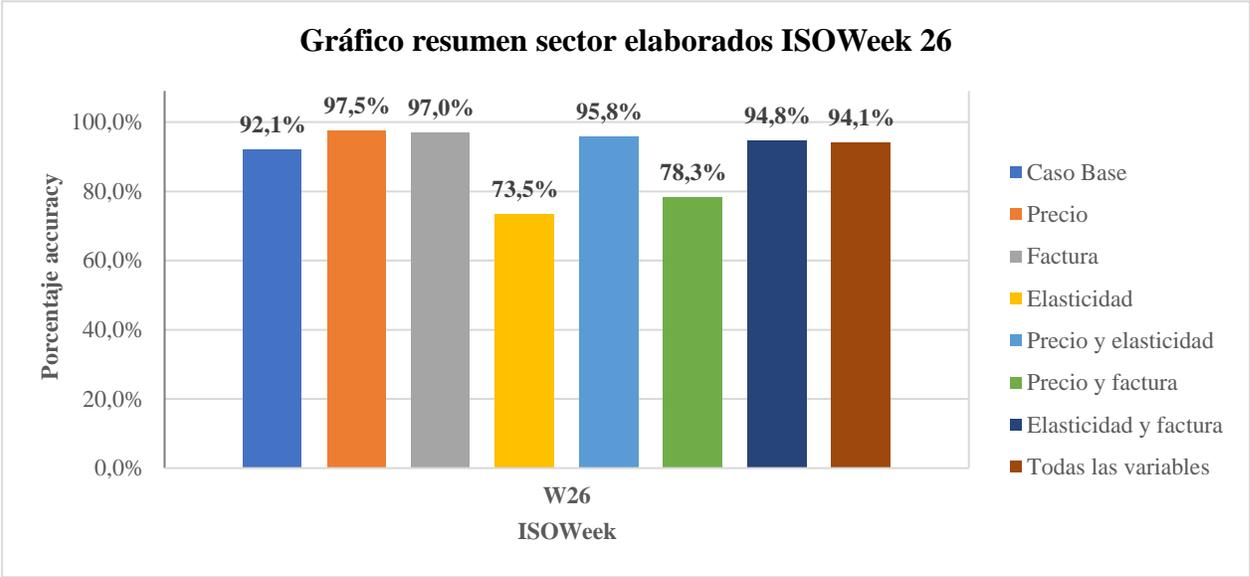


Ilustración 16: Gráfico resumen sector elaborados ISOWeek 26

Fuente: Elaboración propia del estudiante utilizando Excel

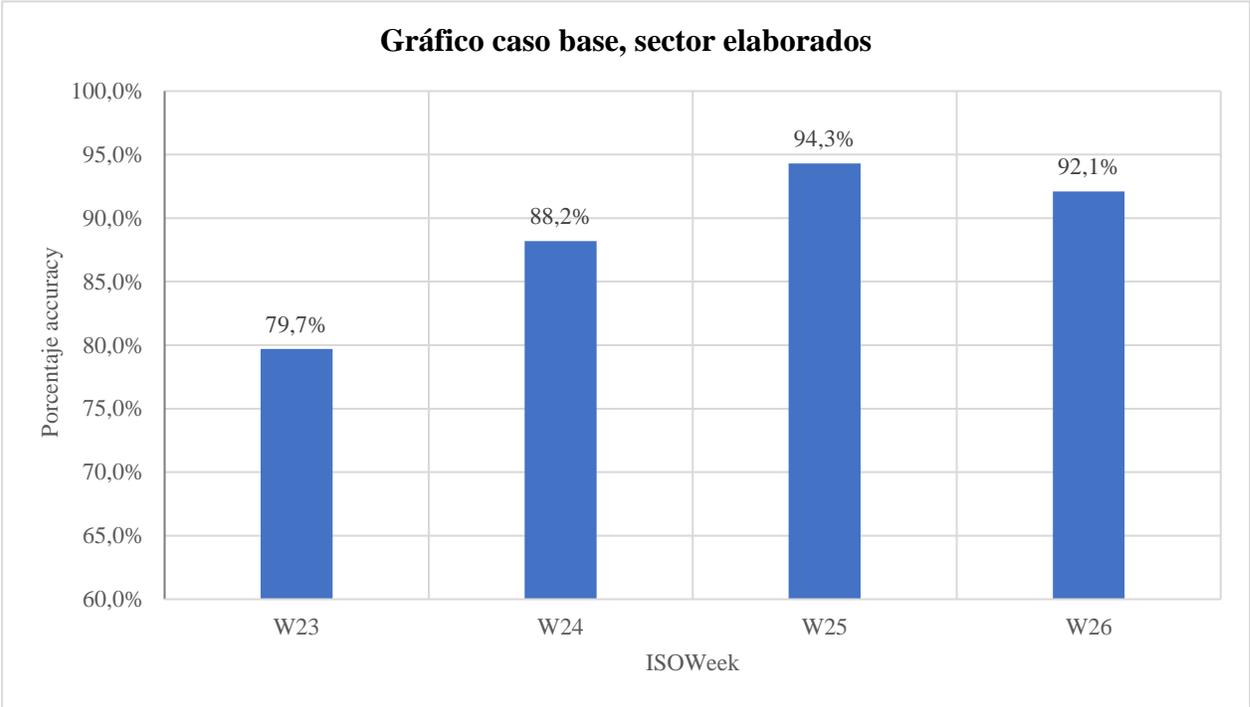


Ilustración 17: Gráfico caso base, sector elaborados

Fuente: Elaboración propia del estudiante utilizando Excel

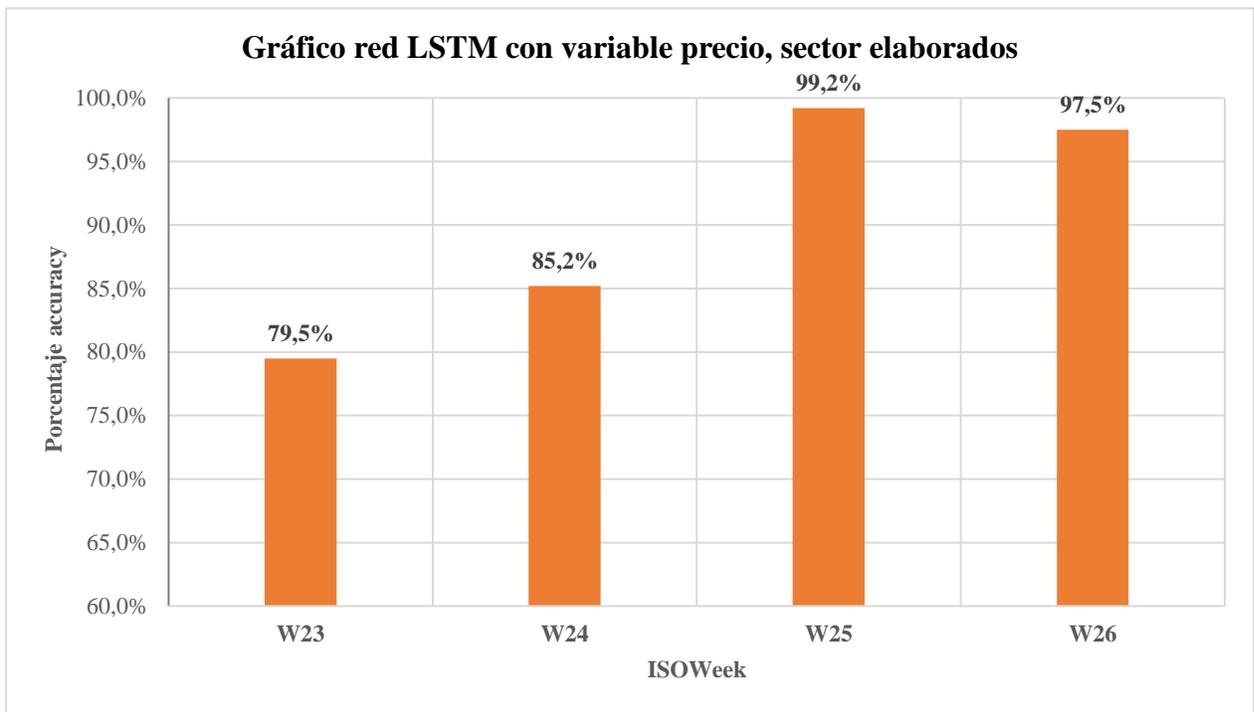


Ilustración 18: Gráfico red LSTM con variable precio, sector elaborados

Fuente: Elaboración propia del estudiante utilizando Excel

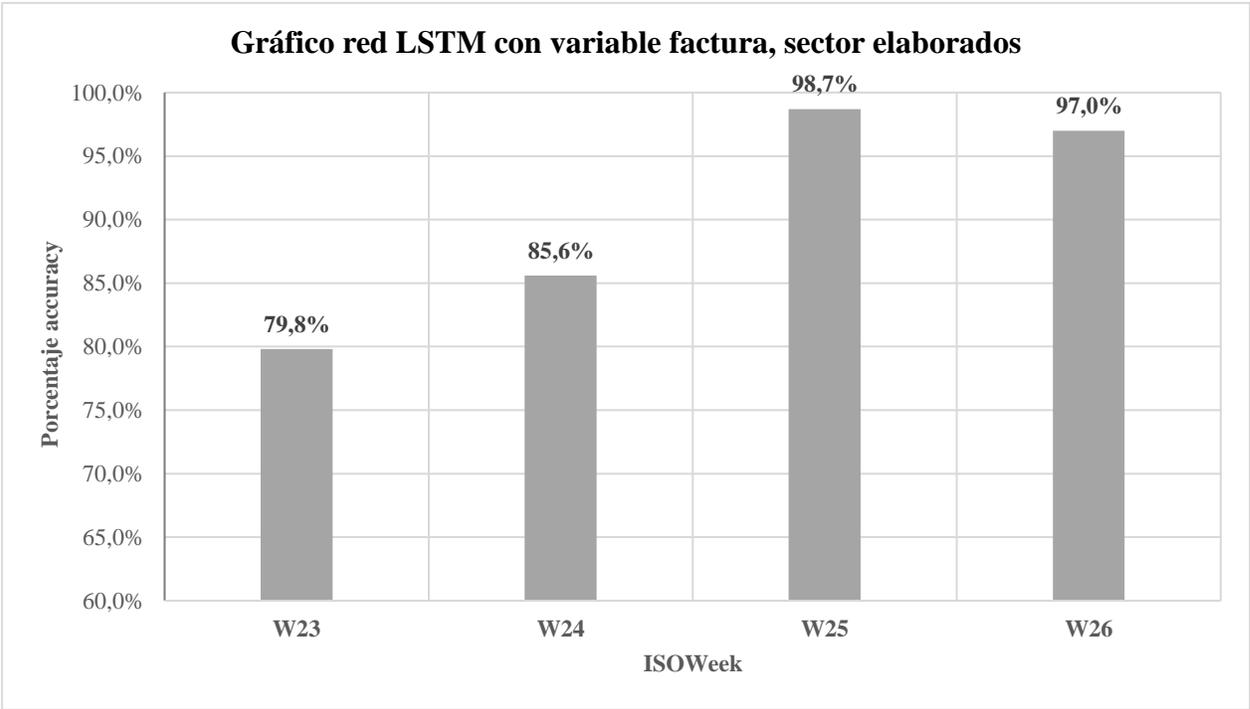


Ilustración 19: Gráfico red LSTM con variable factura, sector elaborados

Fuente: Elaboración propia del estudiante utilizando Excel

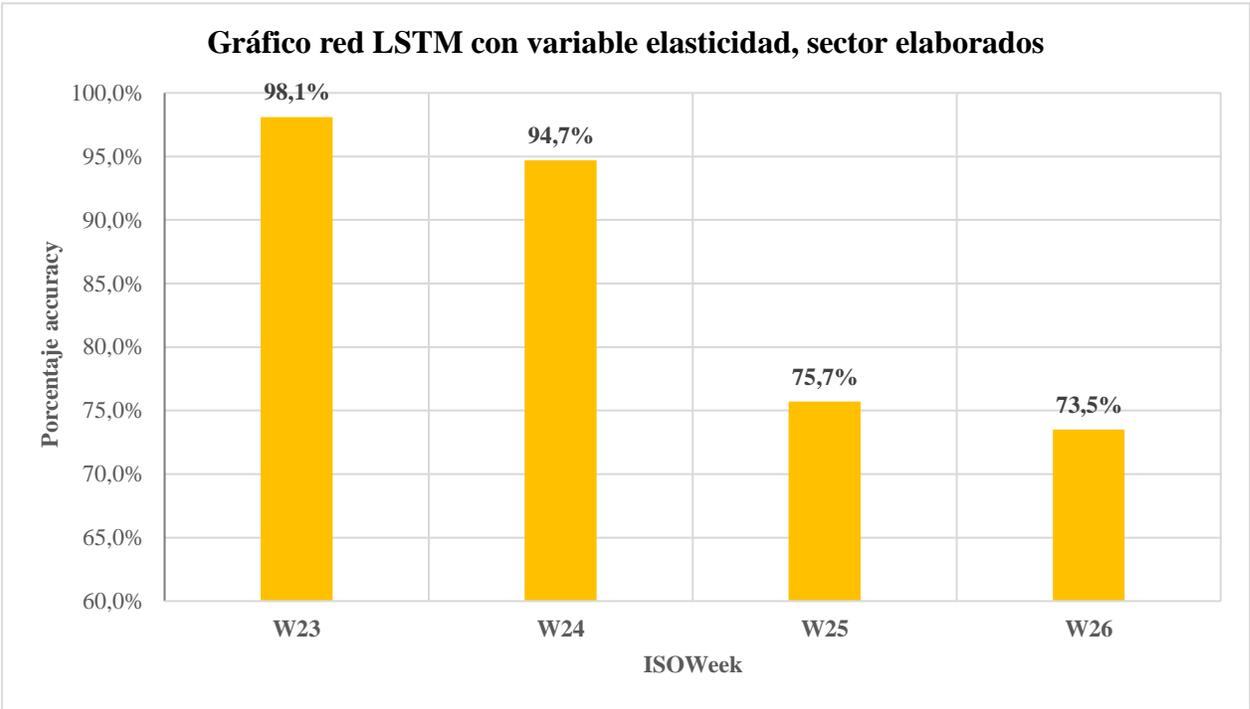


Ilustración 20: Gráfico red LSTM con variable elasticidad, sector elaborados

Fuente: Elaboración propia del estudiante utilizando Excel

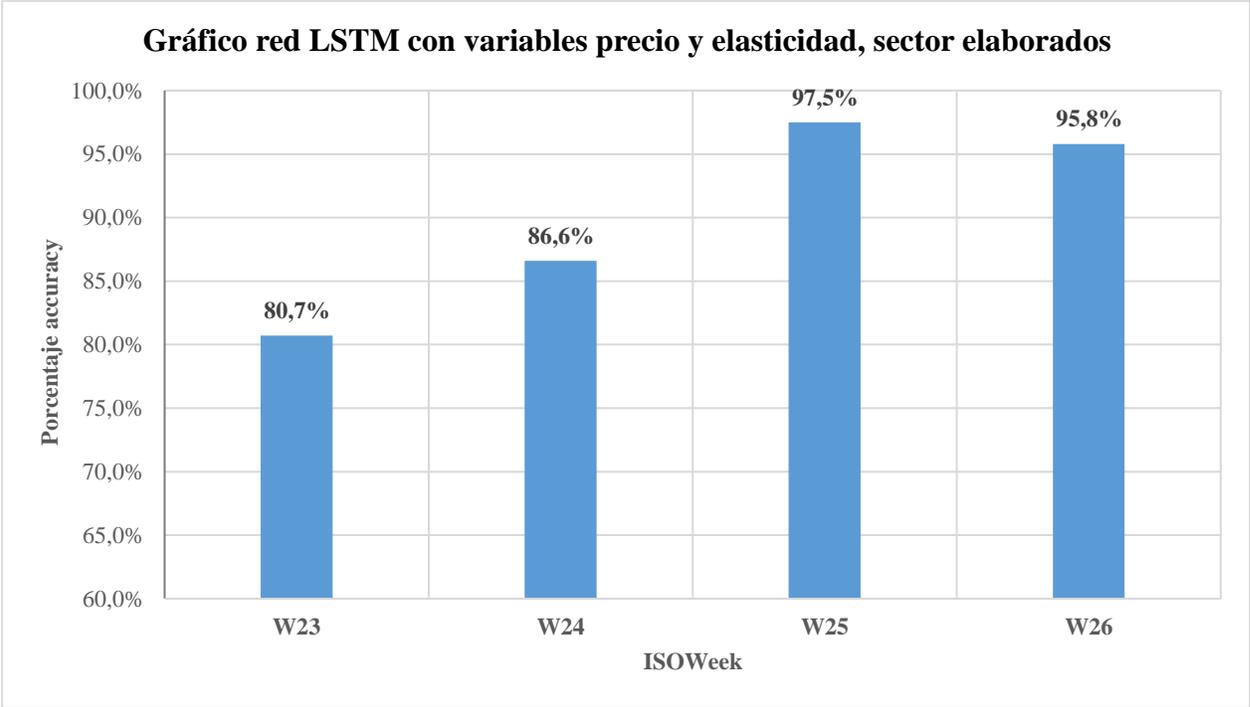


Ilustración 21: Gráfico red LSTM con variables precio y elasticidad, sector elaborados

Fuente: Elaboración propia del estudiante utilizando Excel

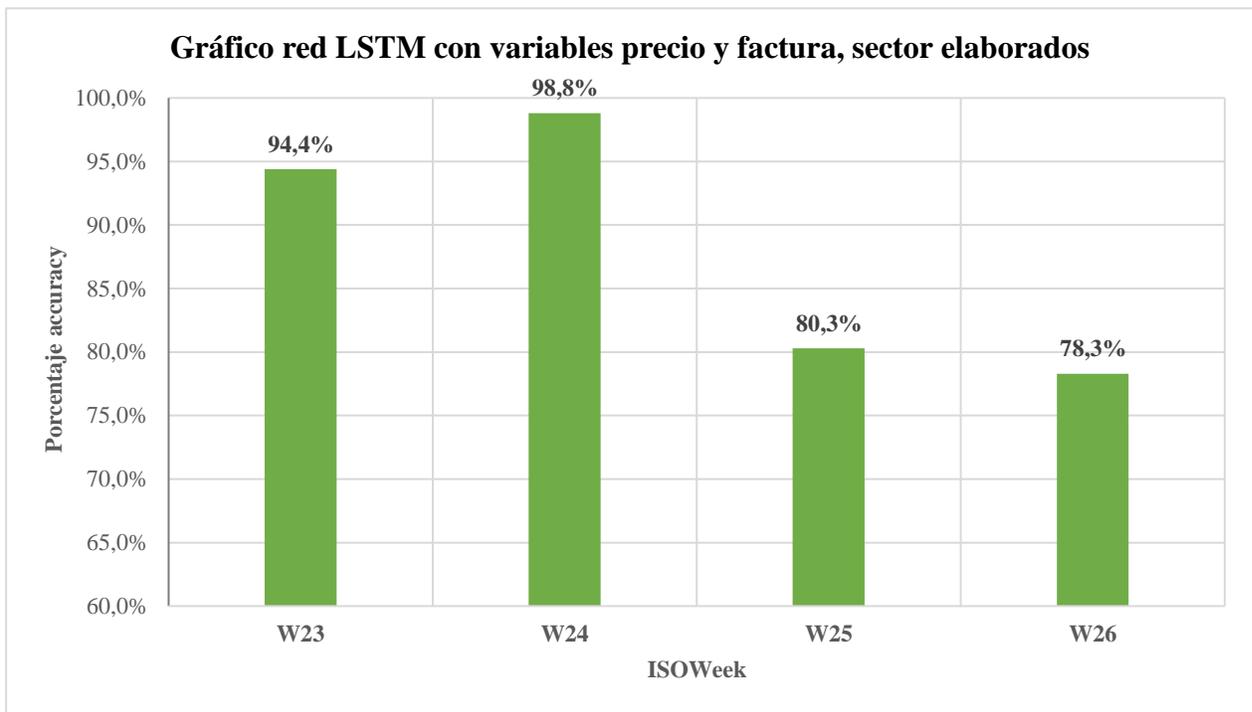


Ilustración 22: Gráfico red LSTM con variables precio y factura, sector elaborados

Fuente: Elaboración propia del estudiante utilizando Excel

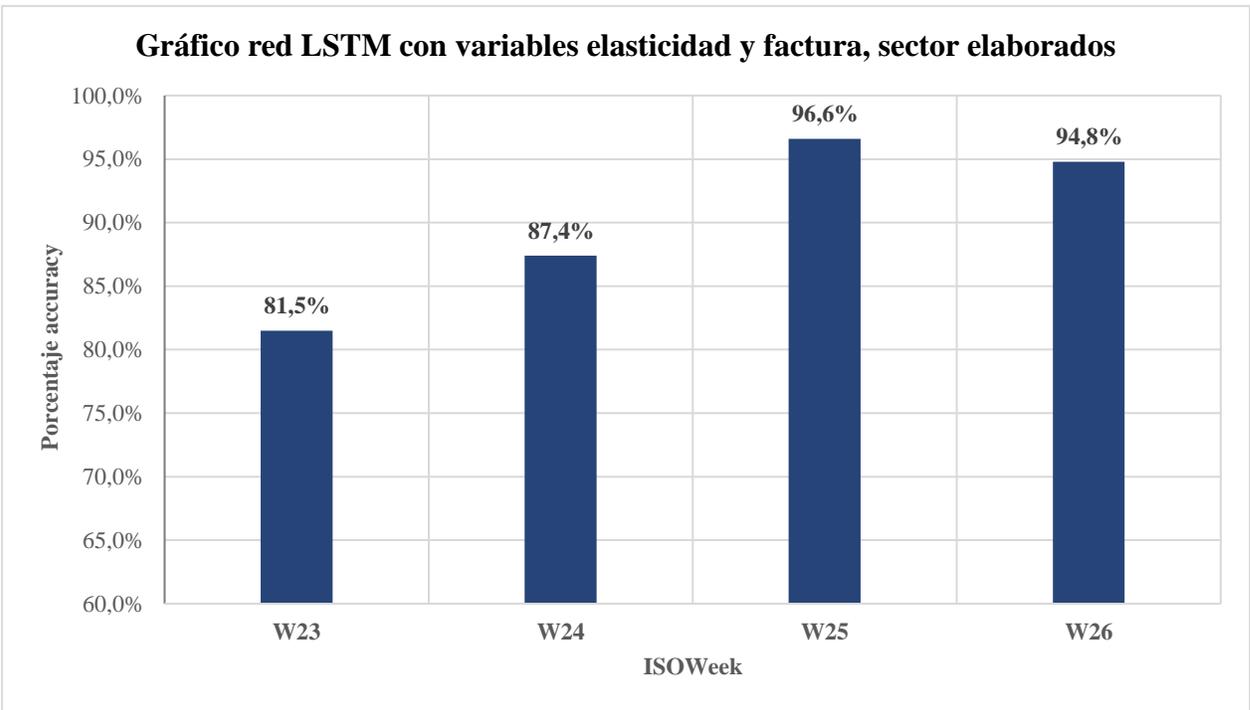


Ilustración 23: Gráfico red LSTM con variables elasticidad y factura, sector elaborados

Fuente: Elaboración propia del estudiante utilizando Excel

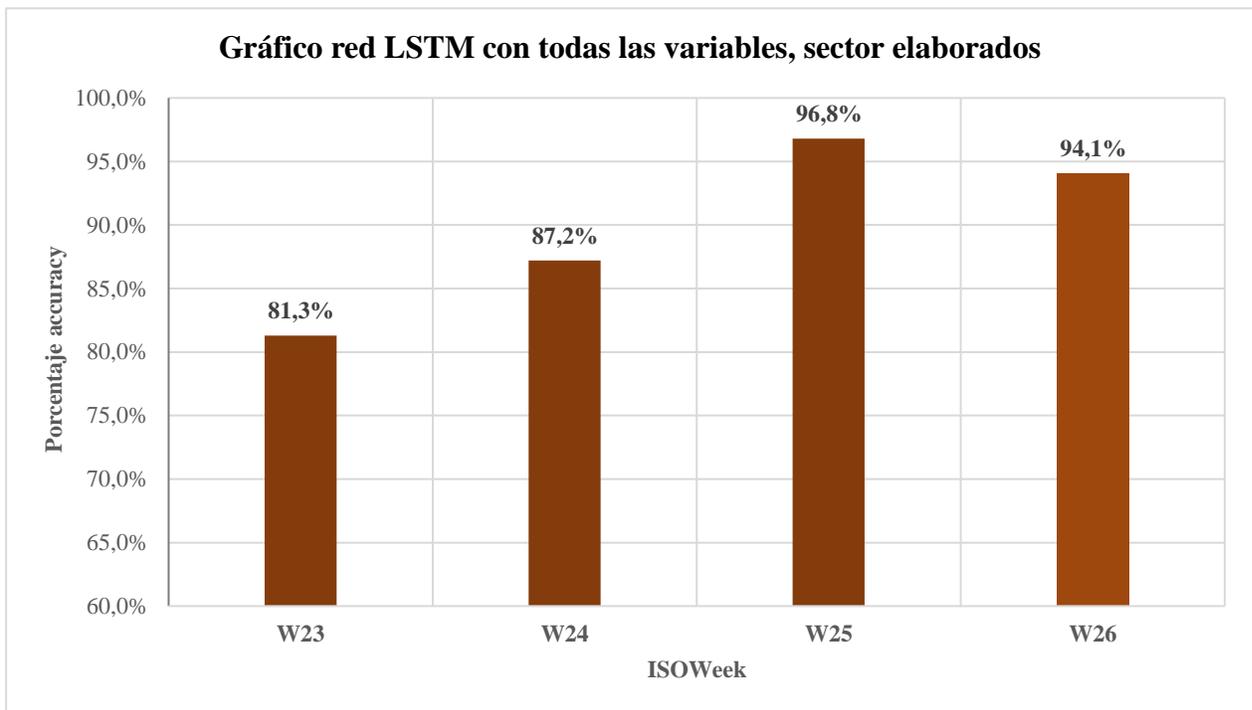


Ilustración 24: Gráfico red LSTM con todas las variables, sector elaborados

Fuente: Elaboración propia del estudiante utilizando Excel

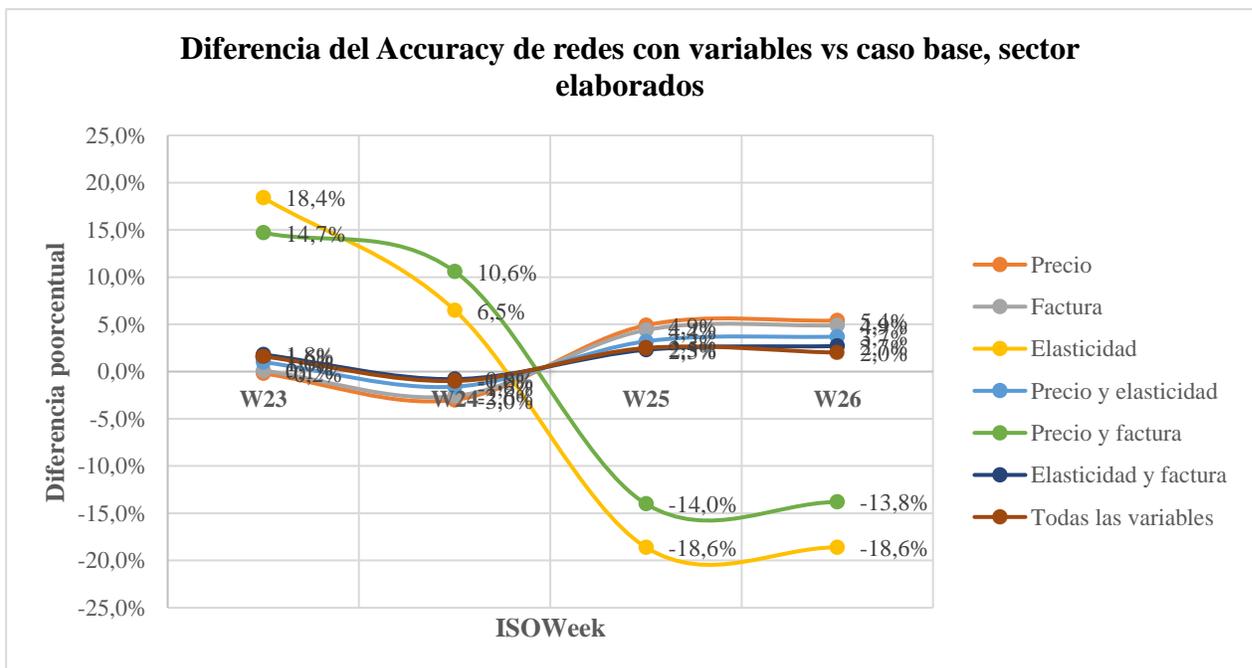


Ilustración 25: Diferencia del Accuracy de redes con variables vs caso base, sector elaborados

Fuente: Elaboración propia del estudiante utilizando Excel

# Sector Cecinas

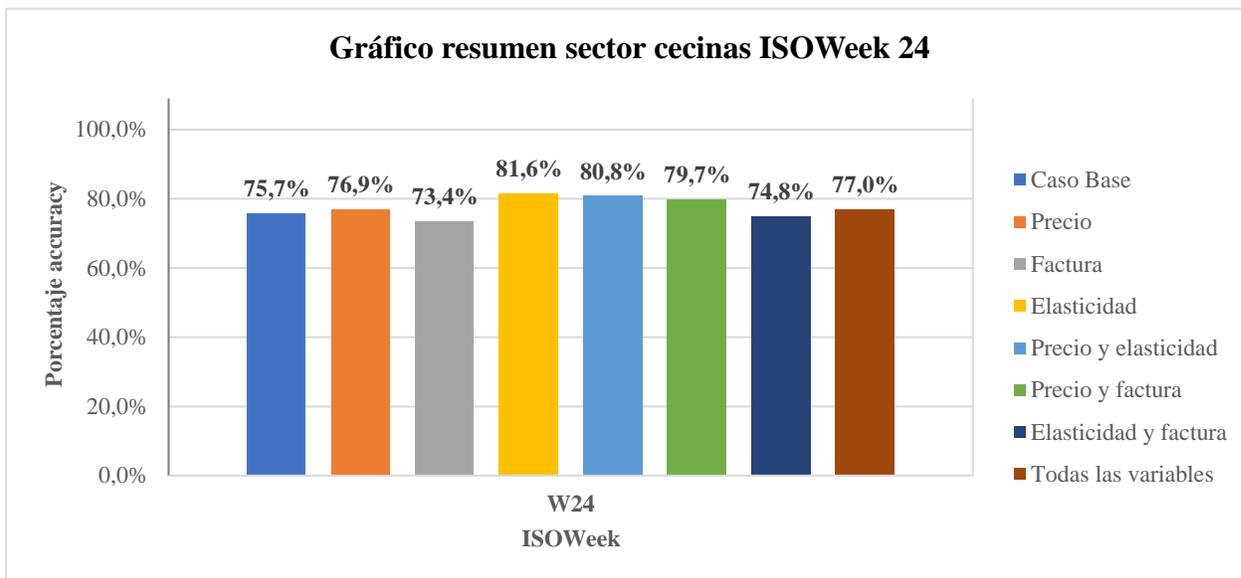


Ilustración 26: Gráfico resumen sector cecinas ISOWeek 24

Fuente: Elaboración propia del estudiante utilizando Excel

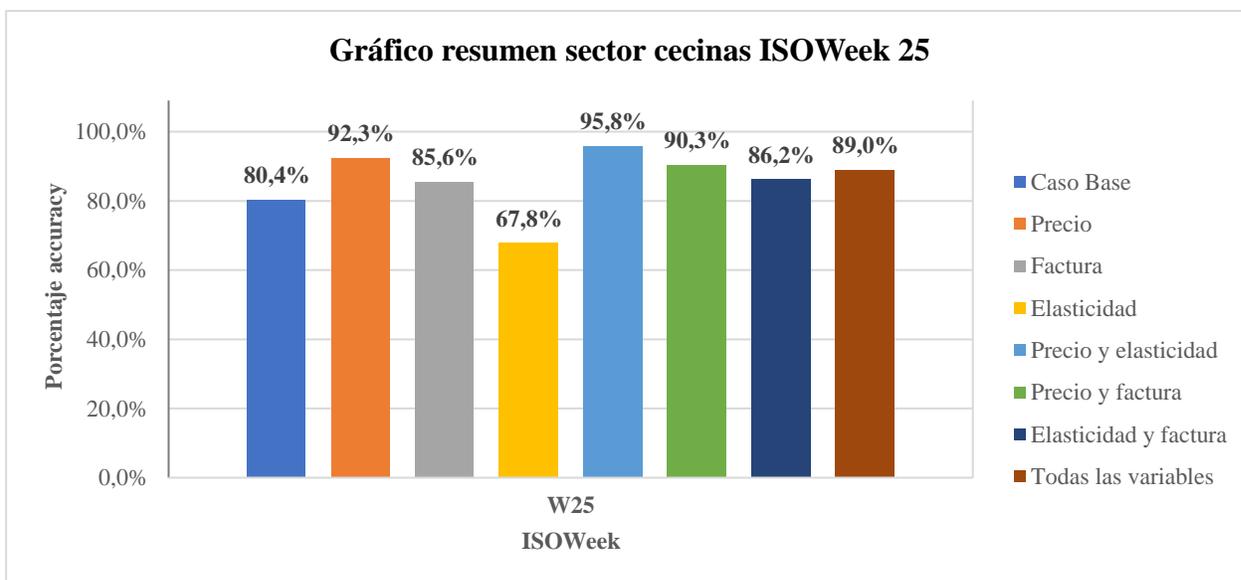


Ilustración 27: Gráfico resumen sector cecinas ISOWeek 25

Fuente: Elaboración propia del estudiante utilizando Excel

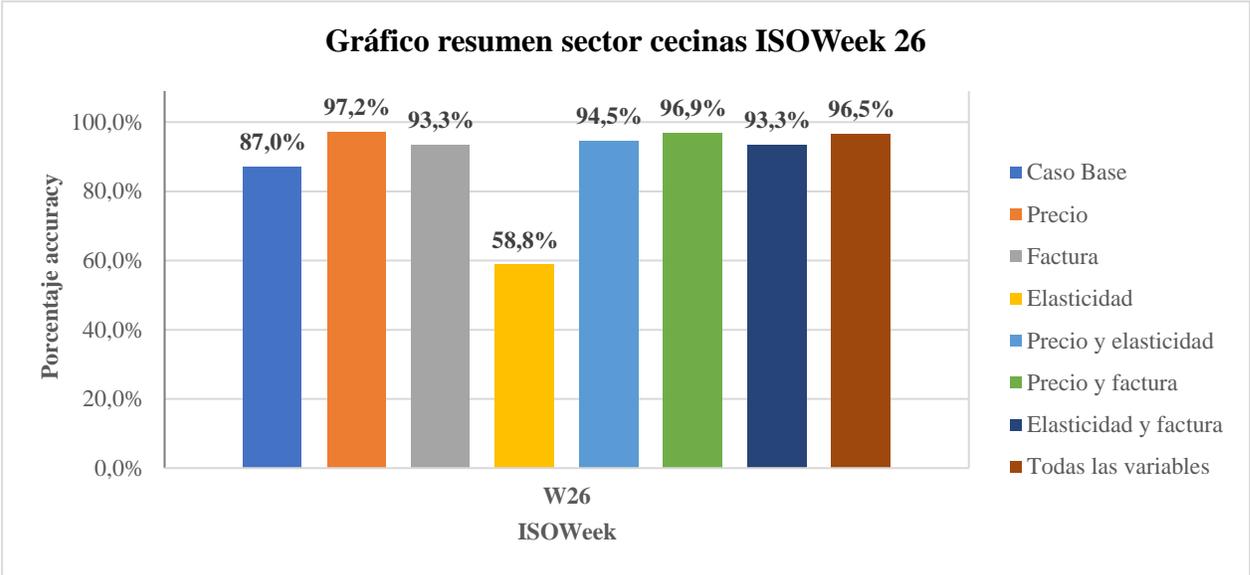


Ilustración 28: Gráfico resumen sector cecinas ISOWeek 26

Fuente: Elaboración propia del estudiante utilizando Excel

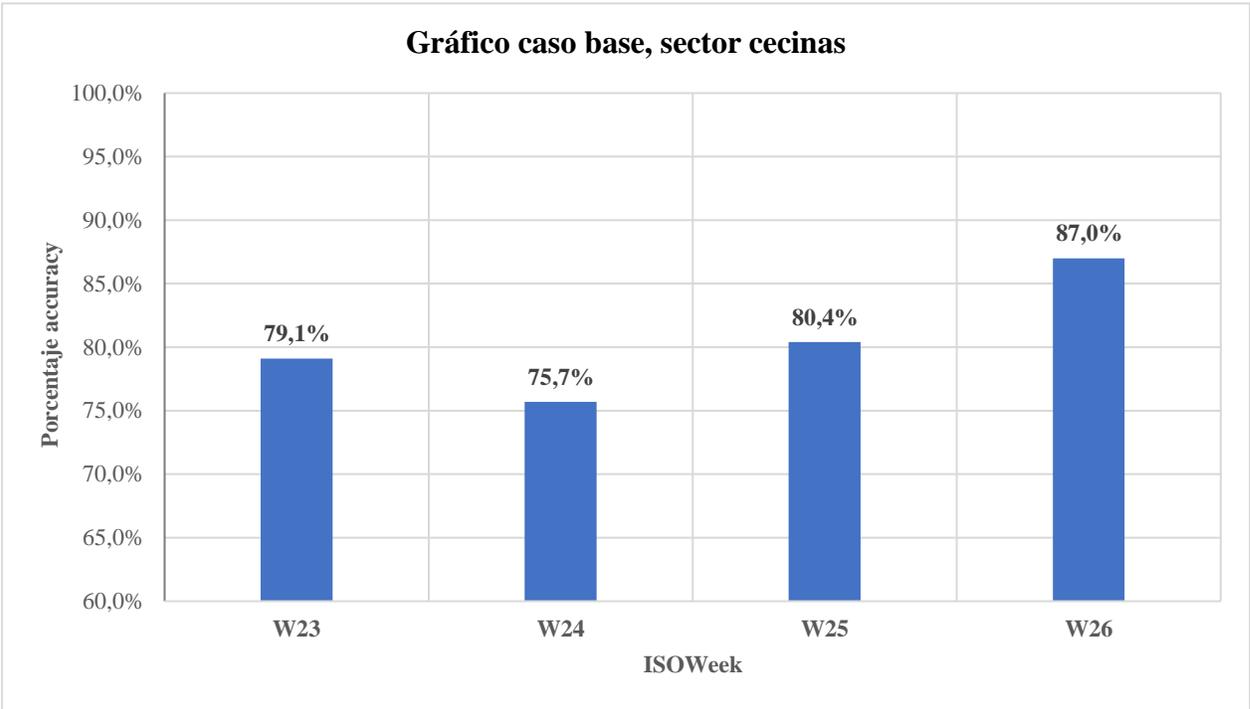


Ilustración 29: Gráfico caso base, sector cecinas

Fuente: Elaboración propia del estudiante utilizando Excel

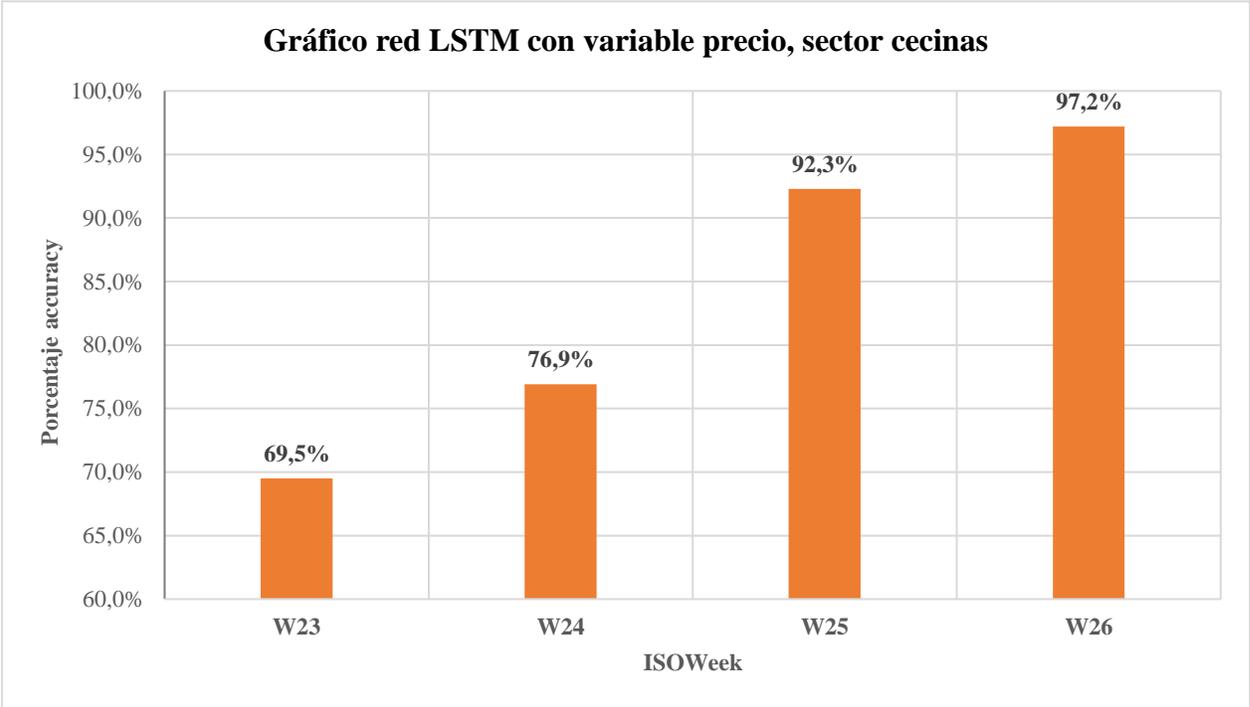


Ilustración 30: Gráfico red LSTM con variable precio, sector cecinas

Fuente: Elaboración propia del estudiante utilizando Excel

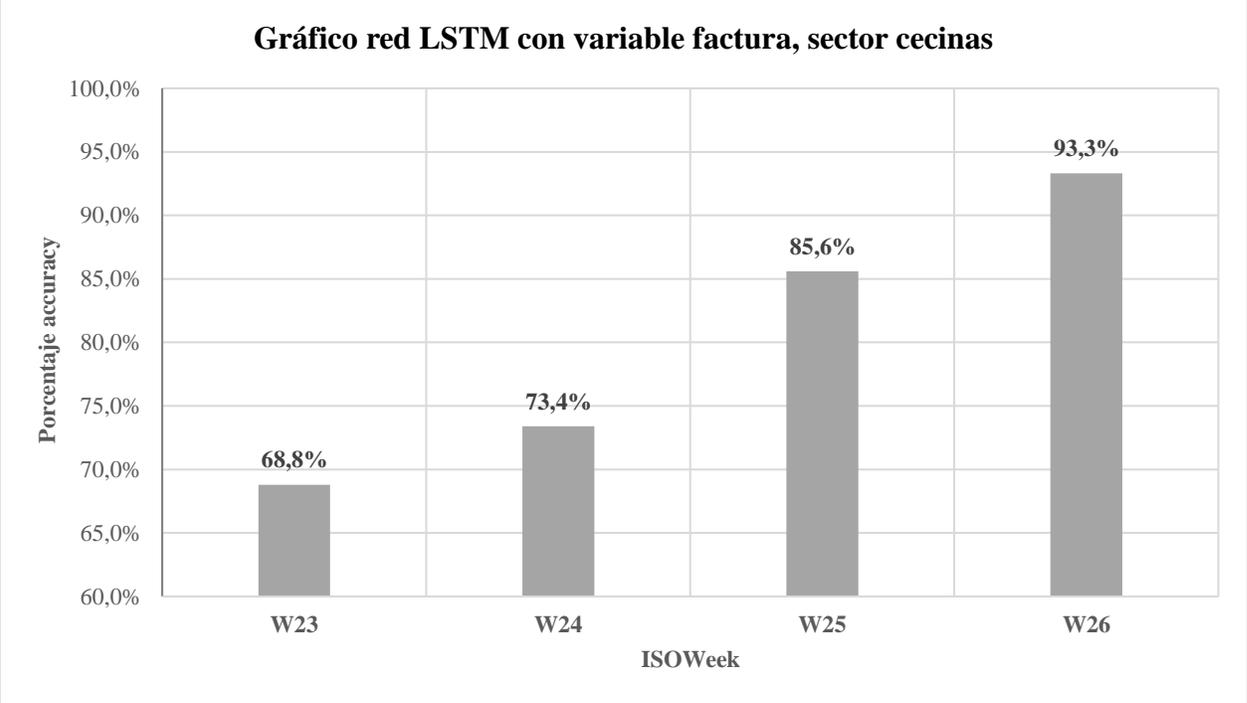


Ilustración 31: Gráfico red LSTM con variable factura, sector cecinas

Fuente: Elaboración propia del estudiante utilizando Excel

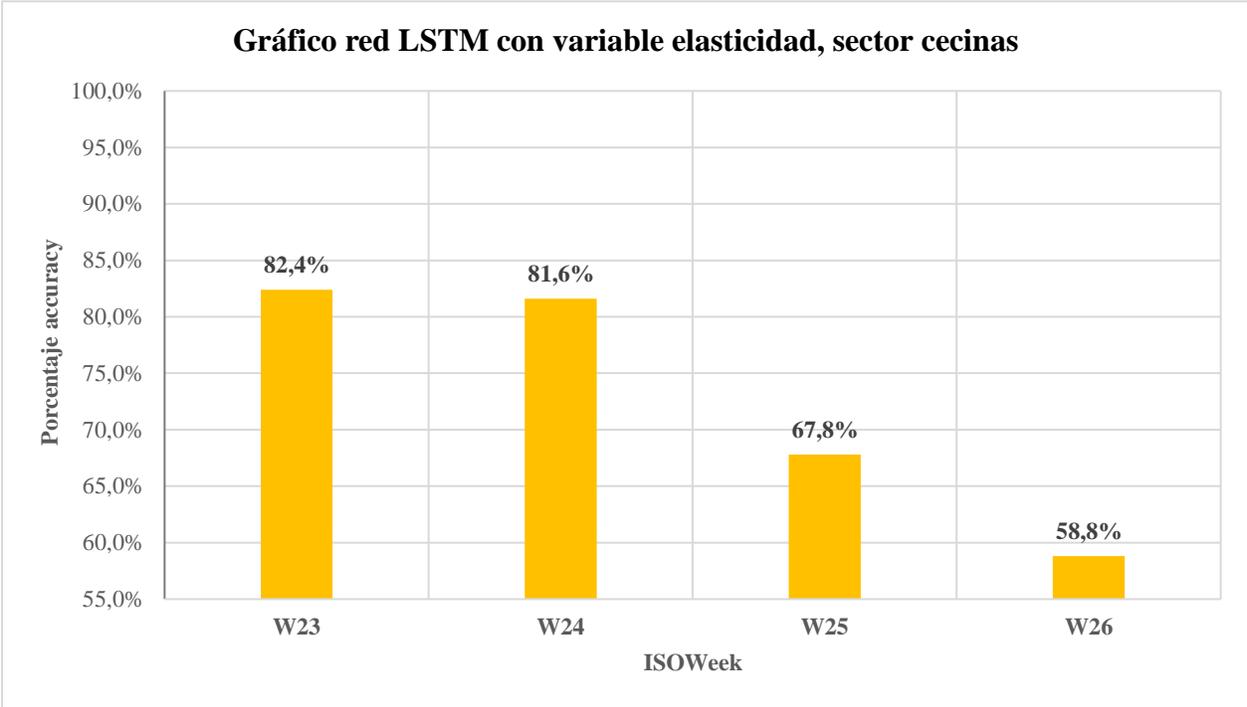


Ilustración 32: Gráfico red LSTM con variable elasticidad, sector cecinas

Fuente: Elaboración propia del estudiante utilizando Excel

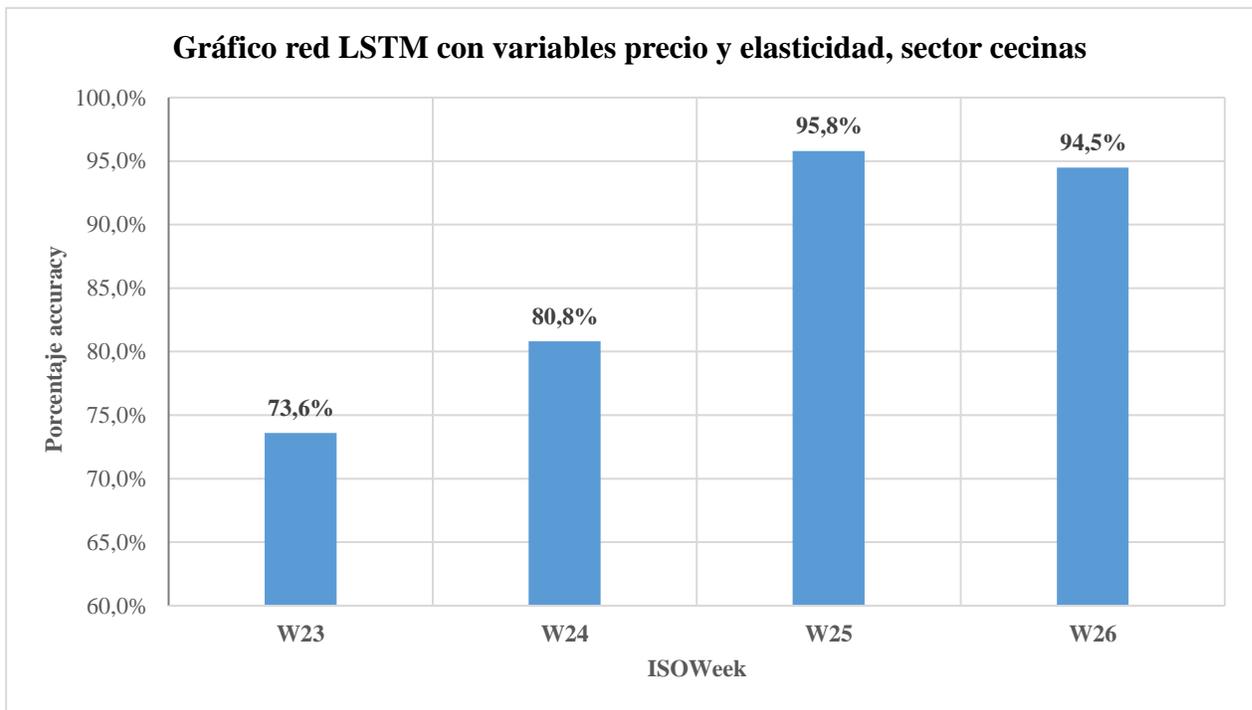


Ilustración 33: Gráfico red LSTM con variables precio y elasticidad, sector cecinas

Fuente: Elaboración propia del estudiante utilizando Excel

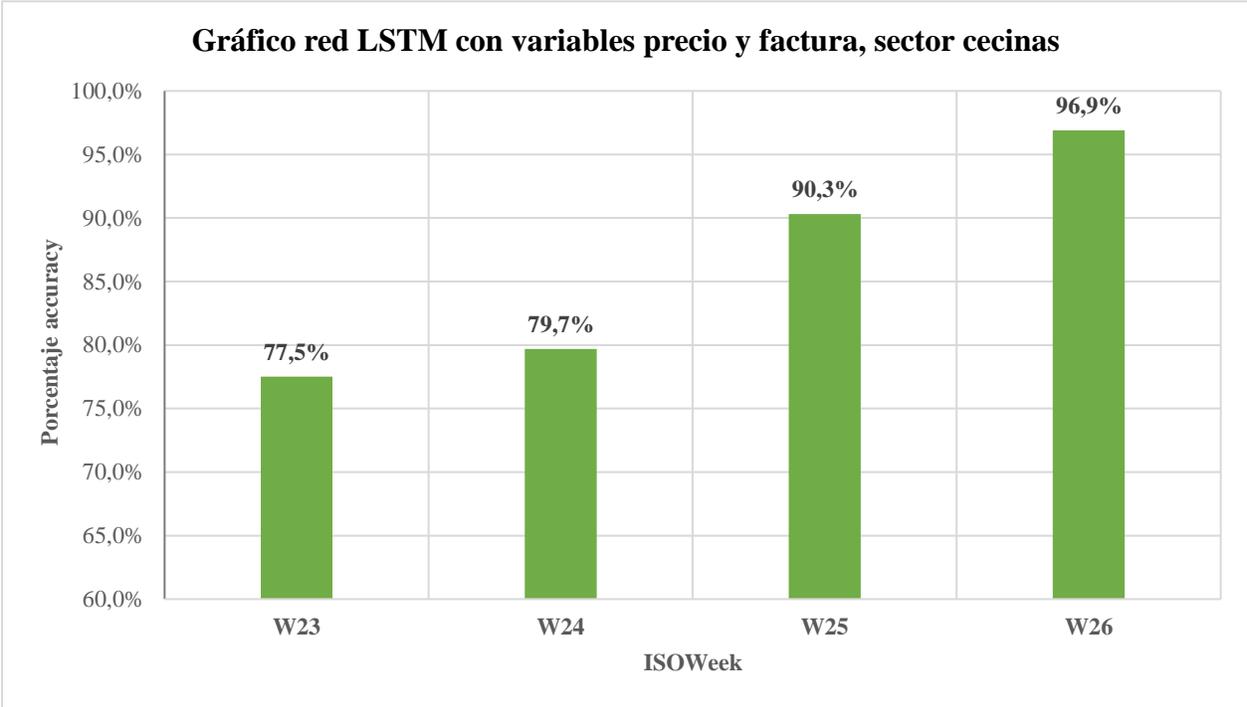


Ilustración 34: Gráfico red LSTM con variables precio y factura, sector cecinas

Fuente: Elaboración propia del estudiante utilizando Excel

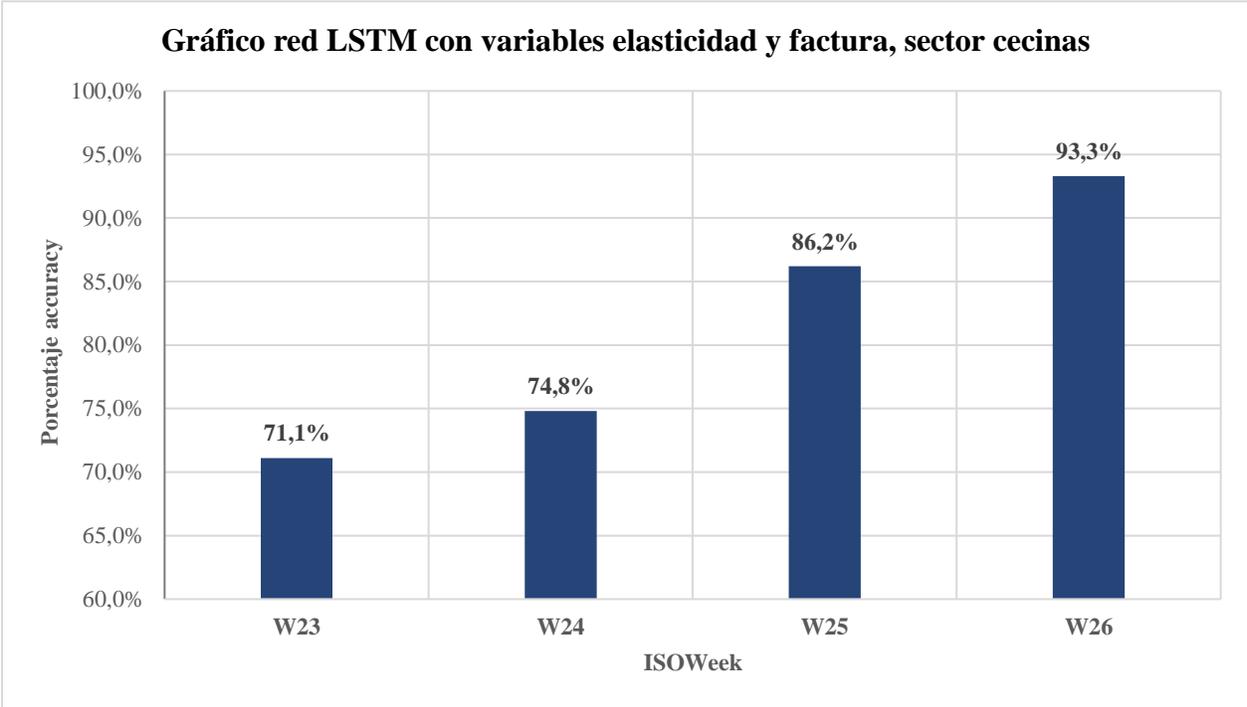


Ilustración 35: Gráfico red LSTM con variables elasticidad y factura, sector cecinas

Fuente: Elaboración propia del estudiante utilizando Excel

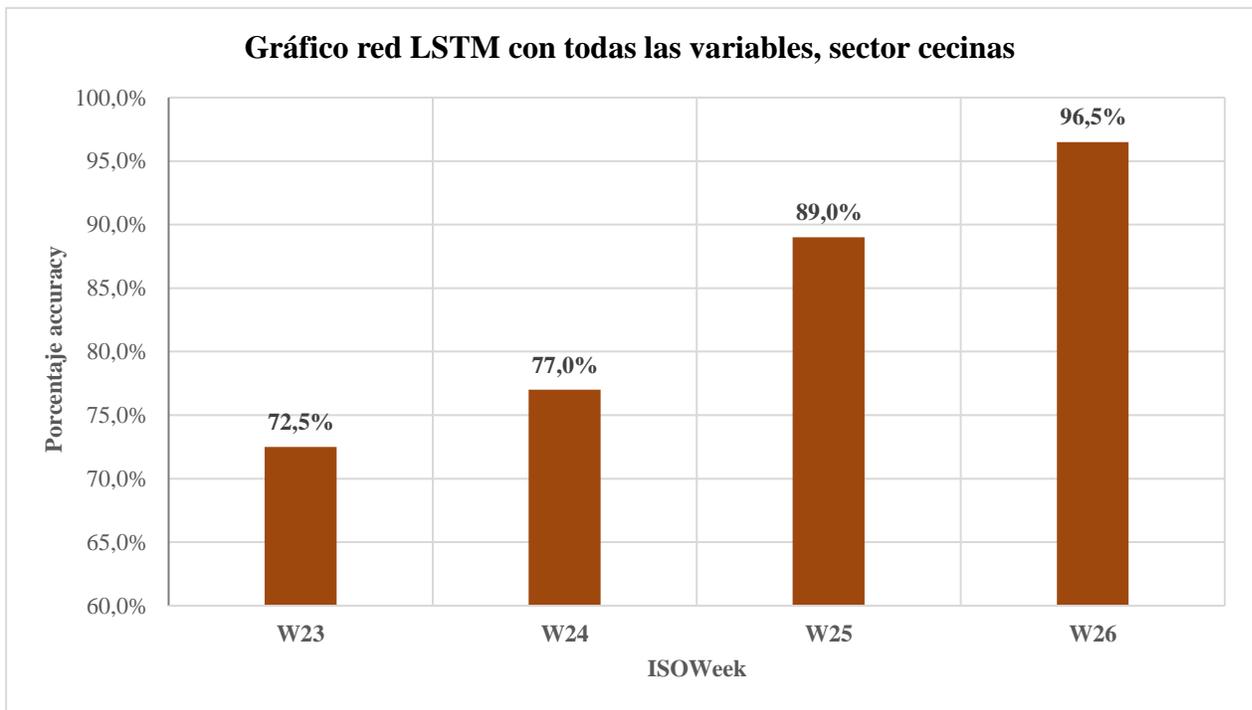


Ilustración 36: Gráfico red LSTM con todas las variables, sector cecinas

Fuente: Elaboración propia del estudiante utilizando Excel

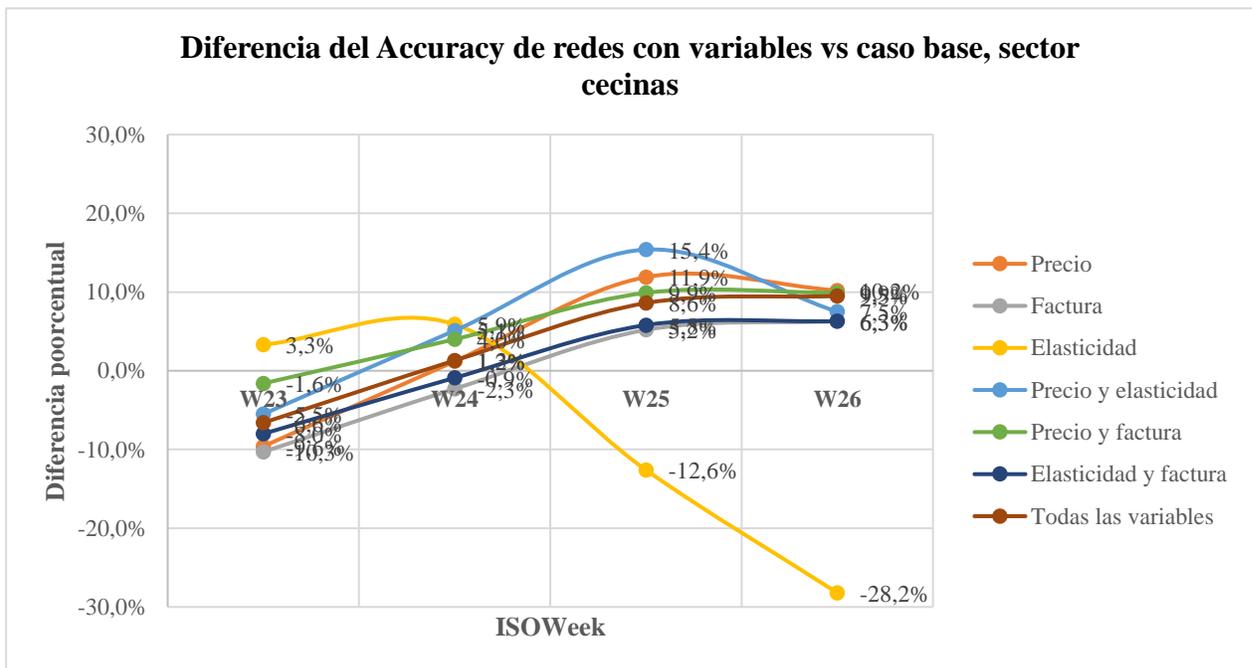


Ilustración 37: Diferencia del Accuracy de redes con variables vs caso base, sector cecinas

Fuente: Elaboración propia del estudiante utilizando Excel

