



**UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL**

**MEJORA DEL PROCESO DE PRONÓSTICO DE DEMANDA EN
PRODUCTOS DE CONSUMO MASIVO EN EL MERCADO NACIONAL DE
AGROSUPER, MEDIANTE UN ENFOQUE DE REDISEÑO DE PROCESOS
Y LA APLICACIÓN DE TÉCNICAS DE DEEP LEARNING**

MEMORIA PARA OPTAR AL TITULO DE INGENIERO CIVIL INDUSTRIAL

IGNACIO ESTEBAN RAMÍREZ RODRÍGUEZ

**PROFESOR GUÍA:
JAVIER SUAZO SÁEZ**

**MIEMBROS DE LA COMISIÓN:
RODOLFO URRUTIA URIBE
FELIPE VILDOSO CASTILLO**

Este trabajo ha sido parcialmente financiado por Agrosuper

**SANTIAGO DE CHILE
AÑO 2022**

RESUMEN DE LA MEMORIA PARA OPTAR AL

TÍTULO DE: Ingeniero Civil Industrial

POR: Ignacio Esteban Ramírez Rodríguez

FECHA: 2022

PROFESOR GUÍA: Javier Suazo S.

MEJORA DEL PROCESO DE PRONÓSTICO DE DEMANDA EN PRODUCTOS DE CONSUMO MASIVO EN EL MERCADO NACIONAL DE AGROSUPER, MEDIANTE UN ENFOQUE DE REDISEÑO DE PROCESOS Y LA APLICACIÓN DE TÉCNICAS DE DEEP LEARNING

El presente trabajo de título se realiza en la subgerencia de Planificación Comercial (PC) de la Cadena de Suministro (CDS) de la empresa Agrosuper, esta compañía es un holding de empresas alimentarias dedicadas a la producción, faena, distribución y comercialización de alimentos cárnicos frescos y congelados de cerdo, aves, salmones y productos procesados. Uno de los aspectos más relevantes para la CDS en temas de gestión de su cadena de valor es la previsión de demanda, la cual toma suma relevancia a la hora de establecer la planificación comercial y de operaciones tanto a nivel táctico(mensual), como operacional(semanal).

La problemática que se aborda en este trabajo son los errores en la predicción de demanda y cómo apuntar a disminuir estos, ya que los niveles de precisión no son los óptimos y según datos de los últimos dos años el coeficiente de variación de la demanda escala en promedio a un valor del 35%. Actualmente, la demanda se pronostica mediante la ejecución del modelo de gestión de S&OP, y según indicadores relevantes (MAPE) que miden la efectividad de esta componente se aprecia que en los últimos meses ha logrado sobrepasar la meta dada por un error de un 18%. La principal hipótesis por validar es si se están ocupando o no los métodos de predicción correctos en el proceso.

El enfoque de este proyecto combina metodologías de rediseño de procesos de negocios y de minería de datos, tiene como principal objetivo mejorar el proceso de estimación de demanda, de manera que esto se traduzca en una optimización de las previsiones de demanda en productos de consumo masivo del mercado nacional. Para esto, se comienza definiendo el marco teórico, se realiza un levantamiento de la situación actual del proceso y de los modelos, para luego proponer mejoras a las prácticas actuales. Además, en base a un plan piloto se mide la efectividad de las proyecciones con métodos de aprendizaje automático (RNA) respecto a métodos tradicionales provistos por SAP IBP, llegando a la conclusión que en el 31% de los casos de previsión del mes de Noviembre se logró un mejor desempeño con este tipo de algoritmo. En particular, se divisa la oportunidad de incorporar esta metodología de predicción en los negocios de carnes de cerdo y de pollo, debido a que estos sectores tuvieron un mejor desempeño en el plan piloto.

Finalmente, se hace una evaluación económica del proyecto, donde se estima que esta oportunidad representa un negocio rentable para la CDS con un VAN de \$25.000.000 aproximadamente, además de entregar una solución innovadora dentro de la industria que puede potencialmente en el largo plazo posicionar a Agrosuper con una ventaja competitiva en el mercado.

Tabla de contenido

1.	INTRODUCCIÓN	1
2.	LA PROBLEMÁTICA	4
3.	LOS OBJETIVOS.....	11
3.1.	OBJETIVO GENERAL.....	11
3.2.	OBJETIVOS ESPECÍFICOS	11
3.3.	ALCANCES.....	11
4.	LA EMPRESA.....	13
5.	MARCO TEÓRICO.....	17
5.1.	CONCEPTOS PRELIMINARES	17
5.2.	MÉTRICAS DE ERROR.....	19
5.3.	MÉTODOS PREDICTIVOS TRADICIONALES.....	21
5.4.	MÉTODO AUTO-ARIMA Y AUTO-SARIMA.....	21
5.5.	MÉTODOS DE ALISAMIENTO EXPONENCIAL	24
5.6.	MÉTODOS DE REGRESIONES LINEALES.....	27
5.7.	MÉTODOS PREDICTIVOS CON REDES NEURONALES ARTIFICIALES	28
6.	MARCO METODOLÓGICO	36
6.1.	METODOLOGÍA CRISP-DM.....	37
6.2.	METODOLOGÍA DE REDISEÑO DE NEGOCIOS DE OSCAR BARROS.....	43
6.3.	PATRONES DE ARQUITECTURA Y DE PROCESOS DE NEGOCIOS	47
6.4.	PATRONES DE NEGOCIO.....	48
6.5.	MODELO Y NOTACIÓN DE PROCESOS DE NEGOCIO (BPMN)	52
6.6.	VARIABLES DE CAMBIO	53
6.7.	METODOLOGÍA FINAL	54
7.	DESARROLLO DEL ALGORITMO DE REDES NEURONALES ARTIFICIALES	56
7.1.	ORIGEN DE DATOS	56
7.2.	SOFTWARE Y LIBRERÍAS UTILIZADAS	56
7.3.	MODELAMIENTO Y DESARROLLO DE ALGORITMOS	58
8.	PROYECTO.....	71
8.1.	CADENA DE SUMINISTRO DE AGROSUPER	71
8.2.	PROCESO DE S&OP Y SU ARQUITECTURA	72
8.3.	LEVANTAMIENTO DE LA SITUACIÓN ACTUAL.....	74
8.4.	SUBPROCESO DE GENERACIÓN DE DEMANDA	78
8.5.	SAP IBP	83
9.	REDISEÑO O MEJORA DEL SUBPROCESO	89

9.1.	DIRECCIONES DE CAMBIO	90
9.2.	TECNOLOGÍAS HABILITANTES	94
10.	RESULTADOS DE PLAN PILOTO	99
11.	PLAN DE IMPLEMENTACIÓN	116
12.	EVALUACIÓN ECONÓMICA DEL PROYECTO	122
13.	CONCLUSIONES Y DISCUSIÓN	126
14.	BIBLIOGRAFÍA	130
15.	ANEXOS	134

1. Introducción

El presente trabajo de título se lleva a cabo en la empresa Agrosuper, buscando aportar en la solución de los problemas de gestión que presenta actualmente, desde la perspectiva general de la Ingeniería Industrial, específicamente desde el ámbito del Rediseño de Procesos de Negocios y la aplicación de técnicas de Deep Learning conocidas como redes neuronales artificiales.

Agrosuper se posiciona en la industria de alimentos, la cual es una de las que tiene una demanda que crece cada vez más en el tiempo y el caso de las carnes no es la excepción. Las evidencias así lo muestran, según estimaciones realizadas por la Organización de las Naciones Unidas para la Alimentación y Agricultura (FAO) la demanda por proteínas de origen animal se duplicará a nivel mundial para mediados de siglo.[1]

En el caso específico de las carnes, la producción mundial deberá aumentar hasta alcanzar los 470 millones de toneladas anuales.[1] Dicho en otras palabras, y en un horizonte más cercano, se estima que su producción escalará en 13% para 2026 respecto del periodo 2014-2016. Aunque es una cifra considerable, es una fracción del aumento ocurrido en la década anterior, el que llegó a casi un 20%. El incremento se logrará, sobre todo, por la mayor producción que se generará en los países en desarrollo. En este escenario, la carne de aves seguirá siendo el principal impulsor en este crecimiento, debido a su menor costo respecto a las carnes rojas. [2]

Según datos de la OCDE, en Chile el año 2018 se tuvo un consumo promedio per cápita de carne de vacuno de 18,7 kilos, mientras que la de cerdo llegó a los 19 kilos y la de aves superó los 36,5 kilos.[3] Esto se condice con la realidad mundial donde las carnes de aves son las que se llevan la mayor proporción del mercado de las proteínas de origen animal.

Según declaraciones del Banco Central (2018) la actividad alimentaria es la segunda en importancia en la economía nacional, con cifras sectoriales que representan el 23% del valor total de las exportaciones; el 20% del valor de las ventas del comercio; el 31% del número de empresas; el 23% del empleo nacional y el 10-12% del PIB, ha tenido un crecimiento del 12,73%, lo que la ubica en uno de los sectores industriales con mejor desempeño.[4]

A la luz de lo anterior, empresas de esta industria como Agrosuper encuentran en el análisis predictivo una herramienta imprescindible para una efectiva gestión de sus operaciones, en específico, se hace relevante lograr tener la mejor estimación y planificación de demanda posible. Al incluir este tipo de análisis en sus procesos les permite, por una parte, sustituir los juicios y apreciaciones personales de los tomadores de decisiones, y, por otra parte, darles capacidad de disminuir las incertidumbres a la hora de elaborar sus estrategias.

La previsión de demanda en Agrosuper se gestiona mediante el modelo de S&OP el que será explicado en este trabajo, pero también se basa en gran medida en juicios obtenidos de personas que trabajan en la cadena de suministro y que cuentan con la experiencia y el conocimiento en este mercado de las carnes, a quienes a menudo llamamos expertos. Se pide a los expertos que predigan lo que podría suceder bajo un conjunto de circunstancias o condiciones cuando no sabemos y es importante estimar qué oportunidades de negocios depara el futuro. Los planificadores comerciales

necesitan encontrar las mejores predicciones sobre la demanda o el precio de los productos que se comercializarán en el futuro. [8]

En el presente trabajo se aplicarán conceptos de la predicción de series temporales, la cual es un área de investigación que ha aumentado su interés últimamente dado el progreso y desarrollo de las Tecnologías de la Información (TI). Este aumento de interés se ve reflejado por la diversidad de sus aplicaciones en diferentes disciplinas en los negocios e industrias, variando desde la economía a la ingeniería, donde la predicción de series de tiempo es una rama de investigación activa y de suma importancia no sólo actualmente, sino que se estima que lo seguirá siendo en el futuro. [9]

El interés en llevar a cabo predicciones más precisas de la demanda futura hace necesario dedicar recursos en investigación para la obtención de herramientas y/o técnicas a la vanguardia y con mejores resultados. Una predicción más precisa en este caso permitirá optimizar la cadena de abastecimiento de Agrosuper y de esta manera tener en stock la cantidad necesaria para vender sin perder ventas ni tampoco almacenar en bodega productos innecesarios. Además, permitirá tomar decisiones sobre la venta y comercialización, en la planificación de la producción, en áreas financieras, entre otras involucradas en estos procesos. Existen muchos ejemplos de series temporales que han sido objeto de análisis para predecir valores u escenarios futuros, principalmente en la ciencia, las finanzas, el comercio e industrias.

El análisis predictivo es una subdisciplina del análisis de datos, utiliza principalmente métodos estadísticos clásicos. Pese a que se evidencian resultados satisfactorios al aplicar estos métodos a series de tiempo lineales, al utilizarlos en series de tiempo no lineales – que, en la realidad, la mayoría son de este tipo – presentan limitaciones porque no son capaces de capturar las relaciones no lineales de los datos. [9]

Frente a lo expuesto previamente, se hace necesario utilizar otras técnicas fuera de las clásicas para realizar predicciones y obtener modelos más eficientes. Las Redes Neuronales Artificiales (RNA) en algunas aplicaciones han mostrado ser una buena opción a la hora de hacer predicciones de series de tiempo, aplicado en un amplio rango de problemas en áreas de comercio, industria y ciencia. [9] Las RNA son una rama de la Inteligencia Artificial (IA) que consiste en el aprendizaje y procesamiento automático, inspirado en la forma funcional del sistema nervioso biológico. A diferencia de los métodos estadísticos clásicos, son capaces de capturar las relaciones lineales y no lineales entre los datos debido a su estructura no lineal que permite un modelo con mayores grados de libertad.

Por lo tanto, predecir científicamente y eficazmente el nivel de consumo de carnes de cerdo y aves a corto, mediano y largo plazo, además de, comprender las tendencias del mercado y la escala general del consumo de carnes son de gran valor para los gestores de la Cadena de Suministro de Agrosuper. Los resultados pueden ayudarlos a planificar la oferta de carnes a nivel nacional, formular políticas de importación y exportación, promover el suministro efectivo de productos y garantizar el equilibrio de la oferta y la demanda de este mercado, e incluso, se piensa que podría ayudar al desarrollo sustentable de la industria de las carnes de cerdo y aves.

Considerando los beneficios que proporciona la correcta estimación y control, pasa a ser una parte fundamental dentro de Agrosuper, impactando directamente en la logística y en los principales procesos de la cadena de suministro. A modo de síntesis, las bondades de un buen pronóstico repercuten de la siguiente manera en la cadena de valor de Agrosuper:

- Disminución de ventas perdidas (gestión comercial y marketing)
- Disminución del stock de seguridad (gestión de stocks)
- Mejora de los términos de negociación con proveedores (gestión de aprovisionamiento)
- Mejoras en la planificación (gestión de producción)
- Optimización en la gestión de pedidos al controlar más la demanda (gestión de pedidos)
- Mejora en el servicio al cliente (servicio al cliente)
- Gestión económica controlada (control económico)
- Compendio de métodos del forecasting

Al existir una variada gama de métodos y técnicas para efectuar un análisis de pronóstico, se distinguen claramente aquellos autores que las identifican como cualitativas y cuantitativas y otros que, además de éstas, incluye a los métodos de proyección histórica haciendo interactuar a los tres métodos con diferentes técnicas a corto, mediano y largo plazo.

En tal contexto se enmarca el presente trabajo de título, que busca generar una mejora en el proceso de estimaciones de demanda en el mercado nacional de Agrosuper, con un enfoque de procesos de negocios, ya que no sólo se trata de un pronóstico que entrega un modelo predictivo en base a los datos históricos, sino que es todo un proceso involucrando a las tecnologías, las personas y los distintos stakeholders del negocio, y un enfoque de minería de datos que busca probar e innovar con la incorporación de un modelo de aprendizaje automático conocido como redes neuronales artificiales, el cual según la literatura es una buena opción a la hora de predecir ciertas variables en otras industrias como la ciencia, la salud, la economía, entre otros campos en los cuales se han aplicado para hacer pronósticos.

2. La problemática

Por lo general, la empresa debido a su experiencia y el conocimiento adquirido a través de su historia ya cuenta con una estructura organizacional, los procesos formalizados y con datos propios para las bases de demanda que incluyen los niveles de servicio para los diferentes segmentos de clientes que involucra esta industria (supermercados, canal tradicional, industriales, entre otros). Además, se tiene en consideración la repercusión de los costos indirectos y las consecuencias de no establecer pronósticos oportunamente.

La principal motivación para realizar este trabajo es que se identificó una oportunidad de mejorar el proceso de pronóstico de demanda y se cuenta con los recursos necesarios para llevar a cabo esa mejora. Existe el acceso a la información y a las tecnologías que apoyen los procesos de selección de datos y minería de datos para descubrir el conocimiento oculto en ellos, ayudando a responder a las preguntas del negocio y tomar decisiones de una manera proactiva basada en conocimiento. Existe la oportunidad de mejorar el proceso integrando la información de fuentes utilizadas actualmente y otras nuevas que se espera que sea un complemento innovador, permitiendo procesarla para generar un conocimiento que sea aplicable a futuros proyectos y que genere una ventaja competitiva con relación a los otros competidores del mercado.

La problemática tratada en este proyecto de título se sitúa en la etapa de pronóstico de demanda, en estricto rigor en la generación de pronósticos estadísticos. En la actualidad el área de planificación comercial no tiene la seguridad si los modelos predictivos están siendo efectivos al momento de predecir y establecer la cantidad de kilogramos de carnes demandados en el mercado nacional, por lo que el foco también está en validar o evidenciar qué tan bien o mal se está llevando a cabo este subproceso de estimación.

Esta problemática es medida a través del indicador de MAPE que representa la diferencia porcentual entre la demanda pronosticada y la demanda real. En los gráficos 6, 7, 8, 9 y 10, se muestra este indicador para cada uno de los negocios de Agrosuper, que representan en parte la efectividad del proceso de planificación de demanda. Se puede comentar que se ha tornado un tanto inexacto, y en cierta medida, se debe a los errores en las predicciones de demanda. Es importante destacar que la meta mensual fijada para el control de este indicador es de un 18%, el cual se muestra en los mismos gráficos.

MAPE Nacional
(Últimos 6 meses)

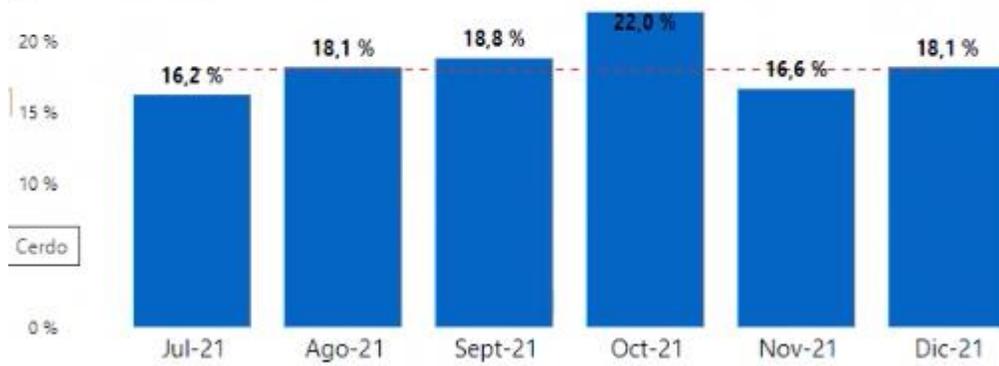


Gráfico 1. MAPE Nacional de los últimos meses del año 2021 para el negocio de pollo.

MAPE Nacional
(Últimos 6 meses)

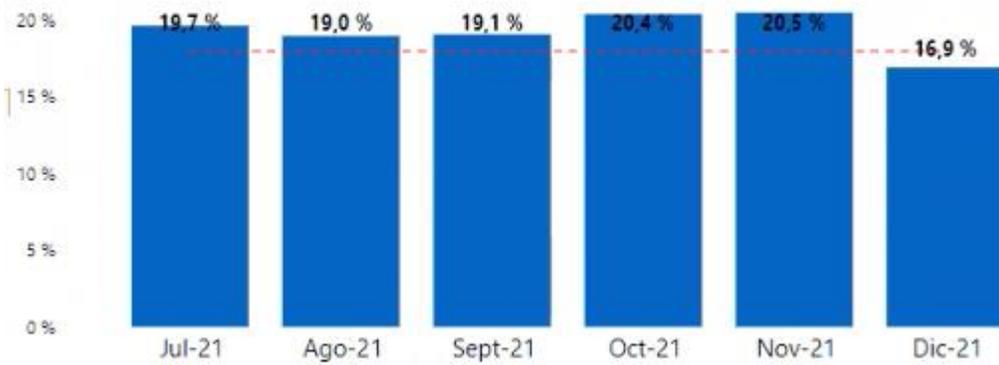


Gráfico 2. MAPE Nacional de los últimos meses del año 2021 para el negocio de cerdo.

MAPE Nacional
(Últimos 6 meses)



Gráfico 3. MAPE Nacional de los últimos meses del año 2021 para el negocio de pavo.

MAPE Nacional
(Últimos 6 meses)

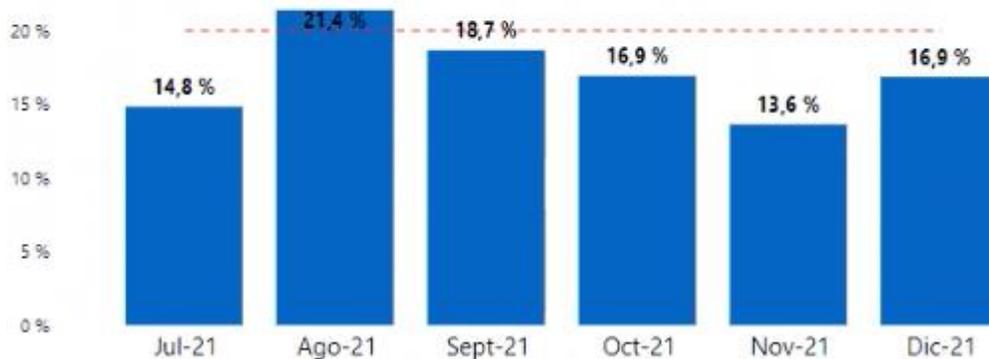


Gráfico 4. MAPE Nacional de los últimos meses del año 2021 para el negocio de cecina.

MAPE Nacional
(Últimos 6 meses)



Gráfico 5. MAPE Nacional de los últimos meses del año 2021 para el negocio de elaborados.

En base a la información levantada para esta problemática se puede comentar lo siguiente:

- El modelo predictivo de demanda que actualmente se utiliza para estos fines es un modelo híbrido que pondera los pronósticos de métodos estadísticos como: Auto-ARIMA, alisamiento exponencial y regresión lineal temporal. Sin embargo, hay una sensación en los colaboradores del área que este modelo ya no está entregando valores muy acordes a lo que la realidad muestra posteriormente.
- No hay un nivel óptimo y sólido en el conocimiento y funcionamiento de estos modelos predictivos y sus principales ventajas/desventajas en su utilización.
- En la etapa de generación de la demanda participan diversas áreas, por lo que existe una alta dependencia de la colaboración de todos, y al momento de integrar las estas perspectivas se producen constantes cambios en los pronósticos.
- En algunos casos las decisiones tomadas y los pronósticos son generados en base a la experiencia, y no necesariamente en base a la información que los modelos predictivos entregan.

- Sólo hace un par de meses se está controlando un indicador de MAPE de demanda, lo cual evidencia que se necesita trabajar y seguir mejorando en este subproceso.
- Los modelos predictivos utilizados para hacer pronósticos de estas series temporales son modelos lineales, en consecuencia, no captan relaciones no lineales que puedan tener los datos, que en general en la realidad la mayoría de las series temporales están afectas a este tipo de relaciones (no lineales).
- Otra causa importante de este problema es que, al momento de predecir la demanda futura, se está tomando como inputs las ventas históricas de productos y no la demanda histórica que son conceptos que deben ser tratados de manera distinta, ya que entre estas dos variables está en juego el nivel de servicio, es decir, puedo tener ventas de X unidades, pero a un nivel de servicio de un 80% por mencionar un ejemplo.

Lo anterior, genera efectos como los siguientes:

- Subestimación y/o sobrestimación de los planes comerciales tanto de manera agregada como de manera desagregada, y, en consecuencia, problemas por conflictos de interés entre ventas y PC.
- El nivel de precisión de los pronósticos de demanda tiene un alto impacto en el resto de la cadena de suministro y de las áreas involucradas, por lo que el no tener una correcta estimación impacta directamente en una ejecución más compleja e ineficiente.
- Volúmenes de inventario indeseados, elevación de costos operacionales y costos de oportunidad, pérdidas de oportunidades de negocio o de ventas.
- Impacto en los niveles de servicio de Agrosuper (fill rate).

El problema del pronóstico de la demanda se presenta hoy en día en casi todas las aplicaciones de las empresas dedicadas a ofrecer bienes o productos; así mismo, es un punto que no se puede obviar en la planificación de la cadena de suministro o en la administración de demanda. Cuando una empresa puede optar por trabajar de acuerdo con la demanda y no de acuerdo con los pronósticos, se puede ufanar de tener un negocio completamente integrado y, por ende, se elimina la necesidad de "adivinar" o suponer la demanda para la elaboración del pronóstico. Al trabajar con predicciones, Agrosuper tendrá una visión de lo que necesitará en el futuro para satisfacer la demanda, además tendrá información de la cantidad y duración de ésta. Con estas predicciones, la empresa es capaz de desarrollar pronósticos de recursos (tiempo, equipos, fuerza de trabajo, compra de partes y materiales) cada vez más precisos.

En el siguiente gráfico se muestran los kilogramos demandados de manera semanal por envasados crudos de los sectores de cerdo, pollo y pavo para tener una dimensión general de las cantidades demandadas:

Demanda por Tipo de Cliente vs Precio promedio

(últimas 7 semanas, por día de la semana)



Gráfico 6. Demandas por tipo de cliente vs precio promedio de productos envasados a nivel agregado de los sectores de cerdo, pollo y pavo. Fuente: Agrosuper.

Por otro lado, es sumamente relevante mencionar que la demanda presenta un comportamiento volátil, por lo que para evidenciar esto se calculó en las siguientes tablas el coeficiente de variación de la demanda de los últimos dos años (2020 y 2021) para los distintos sectores (cerdo, pollo, pavo, cecina y elaborado) y materiales (productos) que Agrosuper comercializa, los cuales serán utilizados para el desarrollo del presente proyecto, se puede observar que estos valores no son pequeños y por tanto se puede inferir que las demandas no son homogéneas durante todos los meses del año, sino que sus valores fluctúan mucho y son heterogéneos entre ellos.

Sector	Material	Desviación estándar (kg)	Demanda promedio (kg)	Coficiente de variación
Cerdo	Chuleta	70.556	485.478	15%
Cerdo	Cost-Pec	70.052	229.388	31%
Cerdo	Cueros	12.284	35.746	34%
Cerdo	Filete	2.676	8.808	30%
Cerdo	Grasa	31.800	125.384	25%
Cerdo	Huesos	13.242	40.708	33%
Cerdo	Lomo	12.113	50.626	24%
Cerdo	Paleta	77.886	447.113	17%
Cerdo	Panceta	2.803	3.762	75%
Cerdo	Pernil	6.345	13.874	46%
Cerdo	Pierna	117.373	509.783	23%
Cerdo	Plancha	9.945	29.720	33%
Cerdo	Prolijado	15.267	62.999	24%
Cerdo	Recortes	27.846	136.686	20%
Cerdo	Reprod	8.511	16.863	50%
Cerdo	Subprod	20.974	75.770	28%

Tabla 1. Variación y promedios mensuales demandados a nivel nacional, para los distintos materiales del sector cerdo de los años 2020 y 2021 de Agrosuper.

Sector	Material	Desviación estándar	Demanda promedio (kg)	Coficiente de variación
Pollo	Ala	46.535	255.216	18%
Pollo	Cazuela	14.579	15.322	95%
Pollo	Con Menudencia	85.729	317.905	27%
Pollo	Filete	39.119	192.966	20%
Pollo	Menudencias	44.787	104.354	43%
Pollo	Patas	3.451	3.698	93%
Pollo	Pechuga	135.380	824.648	16%
Pollo	Pechuga Desech	74.848	413.750	18%
Pollo	Reproductor	9.356	13.582	69%
Pollo	Trutro	283.151	2.125.478	13%
Pollo	Trutro Desechado	4.929	7.275	68%

Tabla 2. Variación y promedios mensuales demandados a nivel nacional, para los distintos materiales del sector pollo de los años 2020 y 2021 de Agrosuper.

Sector	Material	Desviación estándar	Demanda promedio (kg)	Coefficiente de variación
Pavo	Ala	18.419	97.228	19%
Pavo	Menudencias	2.150	7.416	29%
Pavo	Pech Desh	19.697	53.897	37%
Pavo	Trutro	12.387	71.329	17%
Pavo	Trutro Desh	12.185	62.915	19%

Tabla 3. Variación y promedios mensuales demandados a nivel nacional, para los distintos materiales del sector pavo de los años 2020 y 2021 de Agrosuper.

Sector	Material	Desviación estándar	Demanda promedio (kg)	Coefficiente de variación
Cecina	Arrollado	1.511	4.737	32%
Cecina	Fiambre	11.268	45.223	25%
Cecina	Jamón	22.195	132.172	17%
Cecina	Mortadela	10.187	43.984	23%
Cecina	Parrilleros	23.877	65.542	36%
Cecina	Pate	343	668	51%
Cecina	Pechuga	32.842	159.334	21%
Cecina	Prod. Espec.	329	607	54%
Cecina	Salame	893	1.458	61%
Cecina	Salchicha	14.285	39.463	36%

Tabla 4. Variación y promedios mensuales demandados a nivel nacional, para los distintos materiales del sector cecina de los años 2020 y 2021 de Agrosuper.

Sector	Material	Desviación estándar	Demanda promedio (kg)	Coefficiente de variación
Elaborado	Empanadas	1.425	2.300	62%
Elaborado	Empanizado	17.864	106.613	17%
Elaborado	Hamburguesa	36.723	191.690	19%
Elaborado	Moldeado	4.026	23.329	17%
Elaborado	Molido	2.979	10.997	27%
Elaborado	Pizza	2.807	4.924	57%

Tabla 5. Variación y promedios mensuales demandados a nivel nacional, para los distintos materiales del sector elaborado de los años 2020 y 2021 de Agrosuper.

3. Los objetivos

3.1. Objetivo General

El objetivo general de este trabajo de título es:

Mejorar el proceso de estimación de demanda, con el fin de optimizar la efectividad de las proyecciones de demanda en categorías de productos de consumo masivo en el mercado nacional de Agrosuper.

3.2. Objetivos Específicos

Los objetivos específicos de este trabajo son:

1. Realizar un levantamiento de la situación actual (AS IS) tanto del área como de los procesos asociados a la estimación de demanda, identificando oportunidades de mejora.
2. Desarrollar un estudio y modelamiento de los métodos estadísticos tradicionales y de algoritmos de redes neuronales artificiales para comparar la precisión entre estos.
3. Formular una propuesta de rediseño (TO BE) analizando mejoras en el proceso de estimación de demanda a través de la implementación parcial e incremental de un piloto y su respectivo plan de implementación.
4. Evaluar económica y operacionalmente los beneficios del rediseño propuesto a través de la mejora en indicadores definidos en el proceso.

3.3. Alcances

El proyecto de título se desarrollará para la subgerencia de planificación comercial de la cadena de suministro de Agrosuper. Los datos para realizar este proyecto serán otorgados por la empresa para fines netamente asociados a este trabajo de título. Se considera que, para efectos de estudio, modelamiento y análisis de los métodos predictivos; datos de series temporales de demanda en el mercado nacional, entre el periodo de Octubre del 2019 hasta Octubre del 2021, y, cabe destacar que tendrán una frecuencia semanal.

El trabajo estará enfocado en el mejoramiento del subproceso de generación de demanda del proceso general de S&OP. Los sectores analizados para la optimización de los pronósticos de demanda serán el sector cerdo y el sector pollo del mercado nacional de carnes proteicas. Además, cabe destacar que se trabajará con categorías de estos sectores de carnes, como lo son, por ejemplo: la pechuga de pollo, el trutro de pollo, el costillar de cerdo, la chuleta de cerdo, por mencionar algunas. Se desprende que este proyecto tiene un alcance operacional dado que se está trabajando en la Cadena de Suministro (CDS) de Agrosuper.

El software y lenguaje de programación a utilizar para el desarrollo y creación del modelo predictivo mediante redes neuronales será a través de las librerías y entornos de trabajo ofrecidos por Python,

en particular, en un entorno otorgado por Google Colaboratory. En paralelo, para el caso de los métodos estadísticos tradicionales se trabajará con las anteriores, y, por otra parte, con las herramientas y entornos utilizados por el área de Planificación Comercial de la CDS. En particular, SAP IBP y BWP, que son las plataformas utilizadas en donde se está realizando el presente trabajo de título.

Es importante mencionar que, para efectos de implementación del proyecto, esta será parcial y se deja una propuesta final a criterio del área y de la contraparte para que evalúen su implementación, ya sea parcial, total o incluso escalarlo y continuar dedicando recursos para fines que se estimen convenientes para mejorar el proceso y los modelos predictivos.

Con la información provista por la contraparte se ha buscado establecer cuáles son las principales medidas que permitan cumplir con los objetivos señalados anteriormente, modelar el proceso actual, y el propuesto usando un esquema de negocio simple que permita un rápido entendimiento, un Notebook para el caso del modelo de RNA, y una comparación de las métricas de precisión y del indicador del proyecto que permita evaluar económicamente la propuesta.

4. La empresa

Agrosuper es una sociedad anónima cerrada, cuya misión es “alimentar lo bueno de la vida todos los días”, se creó el año 1955 dedicada a la venta de huevos en Doñihue, en la región de O’Higgins. Con el tiempo expandió su negocio a la crianza, el procesamiento y comercialización de aves de corral, cerdo, salmones y procesados (los que incluyen hamburguesas, nuggets, salchichas, croquetas, pizzas, empanadas, jamones y cecinas, entre otros).

La matriz de Agrosuper posee dos segmentos, el de carnes y el acuícola, a través de los cuales ofrece más de dos mil productos, comercializa sus productos a través de 29 oficinas comerciales en el territorio nacional, 11 oficinas comerciales a nivel internacional y bajo las siguientes diez marcas: Súper Pollo, Súper Cerdo, Sopraval, La Crianza, Súper Salmón, King, Súper Beef, AquaChile, Verlasso y Rainforest.

Sus pilares para sostener el nivel de operación están caracterizados por 3 elementos: una robusta infraestructura y red de distribución; salvaguardar la calidad e inocuidad en los productos; procurar el cuidado y bienestar animal. Su cultura organizacional se describe a partir de las siguientes afirmaciones: hacer las cosas siempre mejor; con espíritu de crecimiento y desarrollo; disfrutando la vida.

La compañía declara que durante el año 2020 tuvo: una dotación de personal de 19.507 colaboradores, 71.348 clientes - entre ellos nacionales e internacionales - oficinas en Chile, México, Estados Unidos, Italia, China y Japón, lo que le permite estar presente en 58 países. Los principales destinos de exportación son: China (24%), Estados Unidos (22,8%), Japón (13,1%), Brasil (5,8%), Rusia (5,1%), Unión Europea (5%) y Corea del Sur (4,6%), entre otros.

El mismo año (2020) Agrosuper tuvo ventas totales por 2.594.061 millones de pesos, su participación de mercado a nivel internacional llegó al 0,4% en el caso del cerdo, a 0,5% en el de las aves y 6,3% en el del salmón atlántico. En tanto, se consolida como líder del mercado nacional de carnes en el año 2020 con una participación de mercado de un 55,7% en pollo, un 52,6% en cerdo y un 61,5% en pavo. Además, A fines de mayo de 2018 compró una empresa acuícola para ampliar sus operaciones en el sector y se situó como el tercer jugador de la industria del salmón de Chile, el segundo productor mundial, por detrás de Noruega. En Chile, para el segmento de carnes Agrosuper posee cinco fábricas de alimento para animales y cinco plantas procesadoras, 4 para el sector de carnes y 1 para el acuícola, las que son abastecidas desde sus propias granjas.

A continuación, se representan gráficamente sus clientes con las ventas nacionales segmentadas por canal de distribución y por negocio:

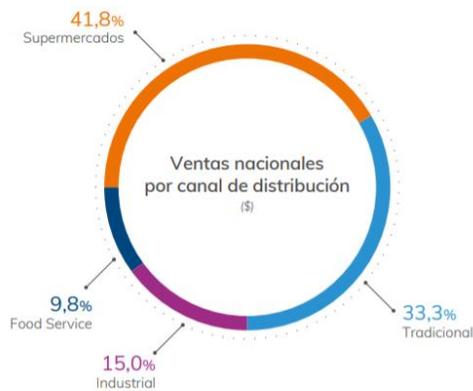


Gráfico 7. Ventas nacionales año 2020 por canal. Fuente: Agrosuper



Gráfico 8. Ventas nacionales año 2020 por negocio. Fuente: Agrosuper

La compañía ha aprovechado las ventajas competitivas de tener su producción en Chile, donde el clima y las barreras naturales permiten una mayor protección y bioseguridad frente a enfermedades, lo que le ha ayudado a desarrollar una imagen de productos de calidad.

Agrosuper cuenta con 7.858 proveedores, declara en su reporte integrado que tienen como objetivo generar relaciones comerciales con personas y empresas que compartan sus estándares y sus normas de integridad. Por ese motivo, buscan crear instancias que fortalezcan la confianza y permitan construir relaciones de largo plazo con sus proveedores. Desde ahí nacen iniciativas como convenio “bolsas de productos” en donde proveedores pueden acceder a financiamiento a tasas menores que el mercado ofrece. Además, según un ranking elaborado por la Bolsa de Productos y la Asociación de Emprendedores de Chile (Asech), Agrosuper destacó en el primer lugar al segmento Acuícola con 86 puntos sobre un promedio de 69,2 puntos de la industria de salmón y al segmento Carnes, en el segundo lugar, con 93 puntos sobre un promedio de 76 puntos de la industria.

Durante estos últimos años la matriz ha avanzado en integrar todos los procesos productivos para asegurar la calidad e inocuidad de cada uno de sus productos llegando a una etapa de consolidación. Para lograr esto, cuentan con un Sistema Integrado de Gestión Agrosuper (SIGAS), basado en los estándares internacionales ISO, IFS y de bienestar animal (OIE), que busca homologar procesos y mejoras, aportar a la sostenibilidad del negocio y cumplir con las expectativas de clientes, consumidores y grupos de interés, mediante un modelo basado en 7 pilares: calidad; inocuidad; seguridad alimentaria; fraude alimentario; bienestar animal; salud y seguridad ocupacional.

La evolución del consumo del segmento de carnes se puede apreciar en los siguientes gráficos:

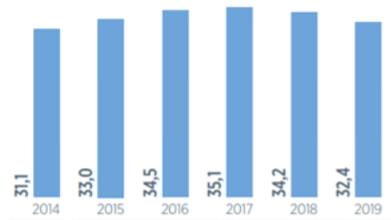


Gráfico 8. Evolución del consumo per cápita del pollo. Fuente: Agrosuper

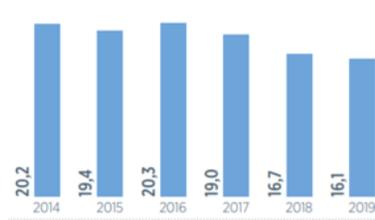


Gráfico 9. Evolución del consumo per cápita del cerdo. Fuente: Agrosuper

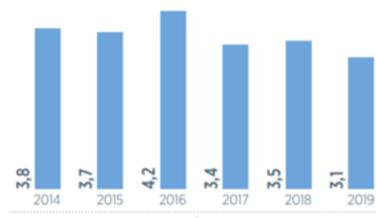


Gráfico 10. Evolución del consumo per cápita del pavo. Fuente: Agrosuper

Su visión como empresa es seguir potenciando durante los próximos años una empresa de alimentos centrada en el consumidor, promoviendo una cultura de colaboración que fomente el desarrollo integral de las personas y del equipo, habilitando el crecimiento del negocio, orgánica e inorgánicamente, a través de marcas y productos de calidad que generen confianza, con preocupación por las inquietudes de los grupos de interés, avanzando en el camino de la transformación digital y maximizando la rentabilidad del negocio con una mirada sustentable de largo plazo.

Ser una empresa de alimentos conectada con el consumidor, ofreciendo productos, soluciones y servicios a través de marcas con propósito que le permitan disfrutar y mejorar su calidad de vida. El negocio de Agrosuper se caracteriza por 4 elementos centrales:

- Crianza animal de pollos, cerdos y pavos
- Faenas
- Distribución/Operación
- Comercialización

Para lograr el éxito de esto último se tienen 5 pilares u objetivos estratégicos:

- Transporte y logística
- Excelencia operacional y mejora continua
- Innovación y transformación digital
- Organización
- Sostenibilidad del negocio

Este trabajo busca ser un aporte que contribuya al tercer pilar específicamente. La estructura organizacional de Agrosuper a nivel macro es del tipo matricial, se puede apreciar en el anexo 1 y 2 de este informe, en dicha matriz se muestra que el segmento carnes está separado del segmento acuícola, pero ambas forman esta compañía. Sin embargo, a continuación, se muestra de manera más específica la estructura organizacional que posee el área donde este proyecto impactará.

5. Marco teórico

Para este apartado es importante comentar que una serie temporal es una sucesión ordenada de datos u observaciones medidos en el tiempo. Una de las utilidades más importantes de las series temporales es su análisis para la predicción de la variable medida, en este caso, la demanda.

Debido al interés creciente por esta tecnología hay una gran cantidad de reportes en la literatura en los que se realizan experimentos sobre distintos conjuntos de datos verificando la performance de las redes neuronales para tareas predictivas. Los estudios coinciden en que los resultados son muy prometedores, pero tal como se indica en [11,12] las conclusiones obtenidas al comparar este tipo de métodos con los estadísticos son inconsistentes entre sí.

De acuerdo con [12], un problema con la literatura asociada a las predicciones de series temporales con aprendizaje automático es que la mayoría de los estudios publicados proveen predicciones aclamando precisiones satisfactorias sin compararlas con métodos estadísticos o inclusive simples puntos de referencia. Este hecho crea altas expectativas con respecto a la capacidad predictiva de las técnicas de aprendizaje automático, pero sin una prueba empírica de que este sea el caso, habilitando conclusiones que podrían no ser sostenibles. Por ejemplo, en [12] se concluye que los métodos basados en Redes Neuronales no son tan buenos como los estadísticos, mientras que en [16,17,18,19] se resuelve lo contrario. Por otro lado, las soluciones que obtuvieron los mejores resultados en las competencias abiertas de comunidades online referentes en ciencia de datos como Makridakis [13,14] surgieron a partir de la aplicación en forma combinada de ambos tipos de métodos como en [15].

5.1. Conceptos Preliminares

Series temporales

Una serie temporal puede ser definida como un conjunto temporalmente ordenado de observaciones que puede ser discreta o continua. Comúnmente tienen una frecuencia fija, lo que significa que las observaciones ocurren en intervalos fijos de tiempo.

El análisis de series temporales es el estudio de series temporales de datos con el objetivo de extraer parámetros relevantes o características de ella. Estos parámetros y características pueden ser luego utilizados para generar un modelo matemático que describa la serie y pueda ser utilizado para realizar predicciones.

De acuerdo con la cantidad de variables o características que la serie temporal contenga se considera univariada para el caso de una variable, o multivariada para el caso de múltiples variables. Dependen de esta fundamental característica los pasos a seguir para su análisis y predicción.

Es importante destacar que una variable de una serie de datos temporales puede no representar solamente valores medidos u observados en el tiempo. Por ejemplo, a partir del procesamiento de una o más variables originales de la serie podría obtenerse una nueva variable aportando información nueva o reforzando información preexistente en la serie [21, 22]. A esto se lo conoce como ingeniería de características.

Estacionariedad

Una serie de datos temporal es estacionaria cuando sus propiedades estadísticas no dependen del tiempo al momento en que la serie es observada. Una forma de ver la condición de estacionariedad es mediante la correlación de sus datos. Si no están correlacionados entre sí entonces la serie es estacionaria. Series temporales con tendencias crecientes o decrecientes, con patrones estacionales o cíclicos, o con varianza variable no son estacionarias. La tendencia y los patrones estacionales afectarán los valores de la serie temporal en momentos diferentes. Por ejemplo, una serie de ruido blanco (señal aleatoria cuyos valores son temporalmente independientes) es estacionaria. La estacionariedad o no estacionariedad de los datos puede ser difícil de verificar a partir de un gráfico de estos únicamente. Una serie temporal con comportamiento cíclico, pero sin tendencia puede ser estacionaria. Esto es así porque los ciclos no tienen una longitud fija, no es posible determinar cuándo serán los picos y valles del ciclo. En general una serie temporal estacionaria no tiene patrones predecibles en el largo plazo.

Es posible convertir una serie temporal en estacionaria por medio de la aplicación de las transformaciones correctas, lo cual es muy importante para el análisis de series temporales, ya que la mayoría de los métodos predictivos estadísticos están diseñados para trabajar con series temporales estacionarias [22]. Esto permite realizar una predicción más precisa y confiable.

Correlación lineal

La correlación lineal entre dos variables es la relación estadística entre ellas. La correlación puede cuantificarse en un valor que, en el caso de ser positivo, significa que ambas variables cambian en la misma dirección.

También puede ser negativo, lo que significa que ambas variables cambian en direcciones opuestas, o, puede ser cero, lo que significa que ambas variables no están relacionadas linealmente [23].

Autocorrelación

La autocorrelación de las observaciones de una serie temporal es la correlación de las observaciones de la serie con las observaciones de una copia de la misma serie desplazada temporalmente. A la función que permite calcular dicha propiedad entre observaciones se le llama función de autocorrelación (ACF, por su sigla en inglés). Representa el gráfico de la cuantificación, en un rango de valores entre 1 y -1, de la autocorrelación de las observaciones. El gráfico de esta función incluye intervalos de confianza establecidos por defecto en un 95% con límites en $\pm 1,96\sqrt{n}$ [23], cuyo significado es que los valores de correlación que estén por fuera de dicho intervalo corresponden a una correlación y no a una anomalía estadística. Esta regla depende de una muestra de tamaño lo suficientemente grande y una varianza finita para el proceso.

Para el análisis de series temporales, la autocorrelación proporciona información sobre cómo están relacionados los datos, las periodicidades y sus frecuencias características. La función ACF de una función periódica tiene la misma periodicidad que el proceso original. A su vez, si las observaciones de una serie temporal no muestran ningún tipo de correlación, significa que dicha serie no es predecible.

Diferenciación

La diferenciación es una de las transformaciones más comunes y convenientes para convertir una serie temporal no estacionaria en estacionaria. Esto se logra computando las diferencias entre observaciones consecutivas. Transformaciones como las logarítmicas pueden ayudar a estabilizar la varianza de una serie temporal. La diferenciación en cambio puede ayudar a estabilizar la media de una serie temporal, por consiguiente, reduciendo la tendencia y los patrones periódicos.

La diferenciación de la serie en términos prácticos es la sustracción del valor actual al valor siguiente. Si Y es el valor de una observación al momento ' t ', entonces la primera diferenciación de la serie puede escribirse como $Y'_t = Y_t - Y_{t-1}$. La serie diferenciada tendrá solamente $t - 1$ valores, ya que no es posible calcular Y'_1 para la primera observación. Si la primera diferenciación no convierte a la serie en estacionaria, es posible realizar una segunda diferenciación. Este procedimiento se puede repetir en caso de ser necesario. En la práctica, sin embargo, no es usual ir más allá de una segunda transformación. La diferenciación afecta a la correlación entre observaciones, por eso si se diferencia de más, se elimina completamente dicha correlación y la serie se vuelve imposible de predecir.

Una manera de determinar de forma cuantitativa si una serie temporal es estacionaria es utilizando tests estadísticos llamados 'Tests de Raíz Unitaria' [24,25]. Son tests de hipótesis estadísticas de estacionariedad utilizados en el análisis de series temporales. Permiten determinar si es necesario aplicarle transformaciones a la serie. Hay múltiples implementaciones de Tests de Raíz Unitaria como:

- Augmented Dickey Fuller test (ADF).
- Kwiatkowski-Phillips-Schmidt-Shin (KPSS).
- Philips Perron test (PP).

Para el análisis del presente trabajo se utiliza la prueba de Kwiatkowski-Phillips-Schmidt-Shin (KPSS) mediante código de Python. En dicha prueba, la hipótesis nula afirma que los datos son estacionarios, luego se busca evidencia para rechazar la hipótesis nula. Por convenio suele establecerse que si el valor de probabilidad devuelto por la prueba es menor al 5% (0.05) entonces es lo suficientemente improbable que lo observado se deba al azar como para rechazar con una seguridad razonable la hipótesis nula. Por eso, para valores de probabilidad menores que 0.05 se asume que la serie no es estacionaria. [24]

5.2. Métricas de error

A la hora de seleccionar una métrica de error es de suma importancia tener en cuenta el campo de estudio, el conjunto de datos y el rol que cumplen los errores. Las distintas métricas de error describen características diferentes de los modelos predictivos y de los datos. Dos de las métricas más utilizadas para problemas de regresión son el error absoluto medio (MAE, por su sigla en inglés) y el error cuadrático medio (MSE, por su sigla en inglés).

MSE

El error cuadrático medio es una métrica simple y común para evaluar un problema de regresión. Se define por medio de la ecuación

$$MSE = \frac{1}{N} \sum (y_t - y'_t)^2$$

donde y_t es la salida esperada e y'_t es la predicción del modelo. Para cada punto, calcula el cuadrado de la diferencia entre las predicciones y el valor real y luego promedia esos valores. Cuanto mayor es el valor de la métrica, peor es el modelo. Nunca es negativo, pero su valor es cero para un modelo perfecto. Es útil cuando los datos poseen valores atípicos, ya que dicha métrica los resalta. Su desventaja es que en el caso de obtener malas predicciones del modelo la elevación cuadrática empeora el problema, sobreestimando lo malo del mismo. Por otro lado, si todos los errores son menores que 1 se obtiene el efecto contrario, se subestima lo malo del modelo. Es utilizado como métrica de la función de pérdida de los modelos de redes neuronales de este trabajo. Cabe destacar que:

$$RMSE = \sqrt{MSE}$$

MAPE

El error porcentual absoluto medio es calculado como el promedio de las diferencias absolutas entre los valores objetivos y las predicciones. Es utilizado en el presente trabajo de título para la evaluación y comparación de los modelos predictivos. MAPE es una puntuación porcentual lineal que significa que todas las diferencias individuales son ponderadas igualmente en el promedio. Por ejemplo, la diferencia entre 10 y 0 será el doble que la diferencia entre 5 y 0. Sin embargo lo mismo no es cierto para MSE. Matemáticamente se calcula utilizando la siguiente ecuación:

$$MAPE = \frac{1}{N} \sum \frac{|y_t - y'_t|}{y_t}$$

MAPE es más robusto, menos sensible a valores atípicos que MSE, pero esto no significa que siempre sea mejor usar MAPE. Por ejemplo, es más indicado utilizar MSE como métrica si es necesario que los valores atípicos tengan mayor preponderancia sobre el modelo predictivo. En caso contrario, de ser necesario disminuir su importancia, es más indicado utilizar MAPE ya que los valores atípicos de los residuales (diferencia entre el valor actual y la estimación del modelo) no contribuyen tanto al error total como al utilizar MSE. Ambas son métricas de error viables, pero describen en forma diferente los errores predictivos de los modelos.

Criterios de información de Akaike, Bayesiano y de Hannan-Quinn

Los criterios de información de Akaike (AIC), Bayesiano (BIC) y de Hannan-Quinn (HQC) son medidas de la calidad relativa de un modelo estadístico, para un conjunto determinado de datos. Proporcionan un medio para la comparación y selección de un modelo entre varios. Dado un conjunto de datos, varios modelos candidatos pueden ser clasificados de acuerdo con su AIC, BIC y HQC siendo el mejor el que presente valores mínimos para cada criterio. Se definen de la siguiente manera:

$$\begin{aligned}
 AIC &= 2k - 2 \ln L \\
 BIC &= k \ln n - 2 \ln L \\
 HQC &= 2k \ln \ln n - 2 \ln L
 \end{aligned}$$

donde en cada caso k es el número de parámetros en el modelo estimado, n es el número de observaciones, y L es el máximo valor de la función de verosimilitud para el modelo estimado. La práctica habitual es ajustar el modelo utilizando el orden de retraso mejor calificado según uno de los criterios anteriores. La librería Statsmodels utilizada más adelante en este trabajo, provee un método que devuelve los valores óptimos de retraso para cada criterio a partir de un modelo y un conjunto de datos.

5.3. Métodos predictivos tradicionales

La idea de estos métodos es creación de un modelo basado en ecuaciones matemáticas que representen la interacción de las variables involucradas.

5.4. Método AUTO-ARIMA y AUTO-SARIMA

Los procesos autorregresivos de medias móviles determinan Y_t en función de su pasado hasta el retardo p , de la innovación contemporánea y el pasado de la innovación hasta el retardo q :

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q}$$

Este modelo se puede escribir en términos del operador de retardos como sigue:

$$(1 - \phi_1 L - \dots - \phi_p L^p) y_t = (1 - \theta_1 L - \dots - \theta_q L^q) a_t$$

$$\phi_p(L) y_t = \theta_q(L) a_t$$

Donde $\phi_p(L)$ es el polinomio autorregresivo y $\theta_q(L)$ es el polinomio de medias móviles.

Un proceso Y_t es integrado de orden d , $Y_t \sim I(d)$, si Y_t no es estacionario, pero su diferencia de orden d , $\Delta^d Y_t$, sigue un proceso $ARMA(p - d, q)$ estacionario e invertible.

El orden de integración del proceso es el número de diferencias que hay que tomar al proceso para conseguir la estacionariedad en media, o lo que es lo mismo, el número de raíces unitarias del proceso. En la práctica, los procesos que surgen más habitualmente en el análisis de las series temporales son los $I(0)$ e $I(1)$, encontrándose los $I(2)$ con mucha menos frecuencia.

En general, si una serie Y_t es integrada de orden d , se puede representar por el siguiente modelo:

$$\Phi_p(L) \Delta^d y_t = \delta + \theta_q(L) a_t$$

donde el polinomio autorregresivo estacionario $\Phi_p(L)$ y el invertible de medias móviles $\theta_q(L)$ no tienen raíces comunes.

El modelo de la ecuación anterior, se denomina modelo Autorregresivo Integrado de Medias Móviles de orden (p,d,q) o ARIMA(p, d, q), donde p es el orden del polinomio autorregresivo estacionario, d es el orden de integración de la serie, es decir, el número de diferencias que hay que tomar a la serie para que sea estacionaria, y q es el orden del polinomio de medias móviles invertible.

El modelo de promedio móvil integrado auto regresivo (ARIMA) es un método de predicción de series de tiempo propuesto por Geogre Box y Gwilym Jenkins [26]. El modelo ARIMA es un método clásico de análisis de series de tiempo y se utiliza ampliamente. El modelo SARIMA se desarrolla sobre la base del modelo ARIMA. Si la serie original tiene una tendencia de fluctuación temporal obvia y características estacionales, se puede considerar el modelo SARIMA [27]. El modelo SARIMA construye un modelo de serie temporal estacional fusionando el modelo ARIMA y el modelo estacional estocástico, que se abrevia como SARIMA (p, d, q) (P, D, Q) S, donde p y q son el orden de la media móvil y autorregresiva respectivamente, P y Q son el orden de la media móvil y autorregresiva estacional, d son los tiempos diferenciales, D son los tiempos diferenciales estacionales, S es el período estacional y la duración del ciclo.

La estacionariedad de las series de tiempo es la premisa básica de las series de tiempo y el procesamiento, y las series de tiempo no estacionarias generalmente deben convertirse en series de estacionariedad. Una vez estabilizada la secuencia por diferencia estacional y no estacional, se determina el orden aproximado del modelo de acuerdo con la función de autocorrelación. Más adelante, se utiliza el criterio de información AIC [28] para determinar la combinación óptima de parámetros del modelo, y el modelo satisface la prueba de ruido blanco residual.

La prueba de ruido blanco es una serie de series normales con distribución independiente, que obedece a la distribución normal con valor medio de 0 y varianza de σ^2 . Cada punto de orden de ruido blanco obedece a la distribución normal. El modelo de ruido blanco se muestra en la ecuación siguiente:

$$x_t = e_t |_{WN(0, \sigma^2)}$$

Pruebe la hipótesis original H_0 : los residuos son series de ruido blanco; Hipótesis alternativa H_1 : los residuos son secuencias sin ruido blanco. Cuando se recibe la hipótesis original, es decir, el residual es ruido blanco, se extrae toda la información del residual y el modelo es bueno.

Para evaluar el modelo de predicción, se prueban el error cuadrático medio (MSE) y el error cuadrático medio (RMSE) entre el valor de ajuste y el valor real. Las dos son métricas de rendimiento que se utilizan comúnmente en el análisis de series de tiempo y se expresan como se mostró previamente.

El modelo cuenta con las siguientes etapas:

- Identificación. Utilizando los datos ordenados cronológicamente se intentará sugerir un modelo ARIMA(p,d,q) que merezca la pena ser investigada. El objetivo es determinar los valores de p, d y q que sean apropiados para reproducir la serie de tiempo. En esta etapa es posible identificar más de un modelo candidato que pueda describir la serie.
- Estimación. Considerando el modelo apropiado para la serie de tiempo se realiza inferencia sobre los parámetros.

- Validación. Se realizan contraste de diagnóstico para validar si el modelo seleccionado se ajusta a los datos, si no es así, escoger el próximo modelo candidato y repetir los pasos anteriores.
- Predicción. Una vez seleccionado el mejor modelo candidato ARIMA(p,d,q) se pueden hacer pronósticos en términos probabilísticos de los valores futuros.

Hay que tener en cuenta ciertas restricciones a la hora de elegir el algoritmo auto-ARIMA/SARIMA ya que funciona más lento que otros algoritmos debido al número de cálculos necesarios y necesita una gran cantidad de datos para lograr resultados efectivos.

Al ejecutar la previsión mediante el algoritmo auto-ARIMA, las tres funciones indicadas anteriormente se combinan en un único algoritmo para pronosticar. La proyección resultante es la suma de los valores revertidos y los errores con el valor medio correspondiente.

Una vez definidos los pedidos máximos, los criterios de información y la estrategia de búsqueda en el modelo de proyección, el algoritmo auto-ARIMA lleva a cabo los siguientes pasos:

1. Comprueba si es necesario efectuar la diferenciación e identifica el pedido óptimo para (d), en caso de que sea preciso.
2. Con el pedido óptimo para (d), calcula las diferencias entre las observaciones consecutivas para transformar los datos no estacionarios en estacionarios.
3. Con los criterios de información seleccionados y el pedido óptimo para (d), calcula distintas combinaciones de (p), (q) y (d) hasta poder identificar el pedido óptimo para (p) y (q). Por ejemplo, una combinación posible de pedidos sería (2,1,3) para (p,d,q).
4. Con los pedidos óptimos para (p) y (q), calcula la proyección aplicando la autorregresión y el valor medio variable a los datos diferenciados.

En el caso de auto-ARIMA estacional (auto-SARIMA), todos los pasos del proceso se llevan a cabo en ciclos estacionales, y no en períodos diferentes. La duración del ciclo estacional la puede proporcionar el usuario o se puede identificar mediante el algoritmo.

Por ejemplo, si se realiza un paso de diferenciación, el sistema sustrae los valores de enero de 2018 (para la diferenciación no estacional) y los valores de febrero de 2018 (para la diferenciación estacional) de los valores de febrero de 2018, siempre que la duración del ciclo estacional sea de 12 meses.

Al aplicar auto-SARIMA, los parámetros relevantes para las distintas partes estacionales se denominan (P), (D) y (Q). Una vez definidos los pedidos máximos, los criterios de información, la estrategia de búsqueda y, de forma opcional, la duración de los ciclos estacionales en el modelo de proyección, el algoritmo Auto-SARIMA lleva a cabo los siguientes pasos:

1. Calcula la duración del ciclo estacional en el caso de que no se haya proporcionado.

2. Comprueba si es necesario efectuar la diferenciación e identifica el pedido óptimo para (d) y (D), en caso de que sea preciso.
3. Con el pedido óptimo para (d) y (D), calcula las diferencias entre las observaciones consecutivas para transformar los datos no estacionarios en estacionarios.
4. Con los criterios de información seleccionados y el pedido óptimo para (d) y (D), calcula distintas combinaciones de (p), (d), (q), (P), (D) y (Q) hasta poder identificar el pedido óptimo para cada parámetro. Por ejemplo, una combinación posible de pedidos sería (2,1,3,0,1,2) para (p,d,q,P,D,Q).
5. Con los pedidos óptimos para (p), (P), (q) y (Q), calcula la previsión aplicando la autorregresión y el valor medio variable a los datos diferenciados.

5.5. Métodos de Alisamiento exponencial

Sencillo

Este algoritmo se puede utilizar para pronosticar la demanda de productos maduros con cifras de ventas bastante estables. El algoritmo detecta fluctuaciones irrelevantes en los datos y los alisa utilizando ponderaciones que disminuyen exponencialmente con el tiempo (es decir, a los datos antiguos se les da paulatinamente una ponderación relativa inferior). La proyección calculada por este algoritmo es un número constante basado en el historial semanal o mensual. Su fórmula para calcularlo es la siguiente:

$$F_t = F_{t-1} + \alpha (D_{t-1} - F_{t-1})$$

Donde F_t es el nuevo pronóstico, F_{t-1} es el pronóstico del periodo anterior α es la constante de suavización y D_{t-1} es la demanda real del periodo anterior.

A diferencia de los métodos de promedio móvil simple y ponderado, este método no necesita de gran volumen de datos históricos de la demanda. Por ende, cada vez que se calcula el pronóstico, se remueve la observación anterior y es reemplazada por la demanda más reciente, y aquí es donde radica la ventaja. Es el método de pronóstico más usado por su simpleza, tanto por pequeñas y grandes empresas, sea en un sencillo archivo de Excel. Entre sus ventajas se pueden mencionar:

- Formulación es sencilla, pues solo requiere el pronóstico anterior, la demanda real del periodo de pronóstico y la constante de suavización, como ya lo veremos más adelante.
- No requiere de gran volumen de datos históricos.
- Es flexible al conseguir darle más importancia a la demanda más reciente o a la antigua.

Su desventaja, al igual que los métodos de promedio móvil, es la respuesta a la tendencia. Aun cuando un valor de alfa (α) logra responder frente a cambios en el promedio, cambios sistemáticos de este harán más grande el error de pronóstico. Es tan así, que cuando se está aplicando un alfa mayor a 0.5 con buenos resultados, optar por el alisado exponencial doble suele ser mejor idea.

Doble

Este algoritmo se utiliza a menudo para crear una proyección para productos en una fase de crecimiento o recesión; es decir, se puede observar una tendencia en los datos. Al aislar la tendencia junto con los valores de ratio, el algoritmo disminuye su efecto en la tendencia de la proyección. La proyección es una línea de tendencia que se basa en el historial semanal y mensual y que es el resultado de una estimación de nivel y de tendencia. Su fórmula para calcularlo es la siguiente:

$$\begin{aligned}F_t &= \alpha D_{t-1} + (1 - \alpha)(F_{t-1} + T_{t-1}) \\T_t &= \delta(F_t - F_{t-1}) + (1 - \delta)T_{t-1} \\FIT_t &= F_t + T_t\end{aligned}$$

Donde

F_t : Pronóstico suavizado exponencialmente con la serie de datos del periodo t

T_t : Tendencia suavizada para el periodo t

D_t : Demanda real del periodo t

α : Constante de suavización para el promedio

δ : Constante de suavizamiento para la tendencia

FIT_t : Pronóstico de demanda con tendencia

El método de suavización exponencial con ajuste a la tendencia requiere de dos constantes de suavización: alfa (α) y delta (δ). Su valor puede estar entre 0 y 1, pero a nivel práctico varía entre 0,05 y 0,50. ¿Cómo escoger los valores más adecuados? Los criterios para definir los valores de las constantes son similares al método de suavización simple.

Para alfa dependerá de la importancia que otorgamos a datos recientes (alfa α más elevada) o a datos más antiguos (alfa α más bajo). El delta funciona similar. Un δ elevado responde con más velocidad a los cambios en la tendencia, mientras que un δ inferior tiende a suavizar la tendencia actual, dando menos peso a los datos recientes.

En la práctica, los valores de α y δ se encuentran con prueba y error utilizando las medidas de error de pronóstico. También se usan softwares.

Triple

Este algoritmo puede utilizarse para modelar una serie cronológica que contenga tanto tendencia como temporalidad. Al aislar la tendencia y la temporalidad junto con los valores de ratio, el algoritmo reduce el efecto que tienen sobre la previsión. El algoritmo necesita al menos dos ciclos de temporada completos de la información del historial de demanda. Su ecuación general es:

$$\begin{aligned}F_t &= \alpha(x_t - E_{t-k}) + (1 - \alpha)(F_{t-1} + T_{t-1}) \\T_t &= \delta(F_t - F_{t-1}) + (1 - \delta)T_{t-1} \\E_t &= \gamma(X_t - F_t) + (1 - \gamma)E_{t-k} \\X_{t+h} &= F_t + hT_t + E_{t-k+h}\end{aligned}$$

Donde:

F_t : Pronóstico suavizado exponencialmente con la serie de datos del periodo t

T_t : Tendencia suavizada para el periodo t

E_{t-k} : Componente estacional donde k es el periodo observado

α : Constante de suavizamiento para el promedio

δ : Constante de suavizamiento para la tendencia

γ : Constante de suavizamiento para la estacionalidad

X_{t+h} : Pronóstico de demanda con tendencia y estacionalidad

De Brown

Este algoritmo es parecido al alisamiento exponencial doble en el hecho de que disminuye el efecto de las tendencias en las proyecciones. En cambio, el alisamiento exponencial lineal de Brown utiliza el mismo coeficiente alfa para alisar el nivel y la tendencia en los datos históricos.

El algoritmo lleva a cabo los siguientes pasos:

- 1- Calcula las diferencias entre los valores alisados exponenciales y los valores alisados dobles.
- 2- Suma las diferencias a los valores alisados.
- 3- Estima los valores de nivel y de tendencia a partir de los valores históricos alisados y suma las estimaciones para cada período.

La proyección resultante responde a modificaciones en la tendencia con mayor rapidez que una proyección calculada por el algoritmo de alisamiento exponencial doble.

Sencillo del porcentaje de respuesta dinámica

Este algoritmo se utiliza a menudo para crear una proyección para productos maduros con números de ventas bastante estables. Es comparable al algoritmo de alisamiento exponencial sencillo salvo en que el sistema recalcula continuamente el coeficiente alfa basándose en el valor anterior de la serie cronológica, lo que puede provocar modificaciones en la proyección ex-post. La proyección es un número constante basado en el historial semanal y mensual.

Automático

Este algoritmo selecciona automáticamente el algoritmo de alisamiento exponencial apropiado e identifica los coeficientes óptimos antes de calcular la previsión. El proceso de selección depende del alcance de optimización que se seleccione. Si el planificador comercial marca "Seleccionar algoritmo de alisamiento con mejor resultado", el algoritmo realiza los pasos siguientes:

1. Verifica si se puede observar algún patrón de tendencia o estacionalidad en los datos históricos.
2. En función de los resultados, elige el algoritmo de alisamiento exponencial triple, doble o simple para el cálculo de pronóstico.
3. Calcula varios pronósticos ex-post con el algoritmo de alisamiento exponencial elegido y varios coeficientes alfa, beta, gama y phi.

4. Compara los pronósticos ex-post calculados con la medida de error seleccionada (MSE o MAPE) y utiliza los resultados para elegir los coeficientes óptimos.
5. Calcula la previsión con el algoritmo de alisamiento exponencial y los coeficientes seleccionados.

Cabe destacar que en el área de Planificación Comercial de Agrosuper este algoritmo es utilizado para efectos de pronósticos con este tipo de modelo predictivo.

5.6. Métodos de regresiones lineales

Temporal

Este algoritmo identifica una función lineal para los datos de ventas históricos, en función de los cuales se calcula la proyección. Mientras que la regresión lineal puede prever tendencias, la regresión lineal estacional también puede tener en cuenta la temporalidad si es necesario. La ecuación de este método se expresa como sigue:

$$Y = \alpha + \beta X + e$$

Donde α representa la intersección con el eje y , β representa la pendiente de la línea ajustada, y e representa el error asociado.

Ocasionalmente, este algoritmo podría producir resultados más sólidos que los algoritmos alisamiento exponencial triple y aislamiento exponencial automático, que también se pueden utilizar cuando las tendencias o la temporalidad se pueden comprobar en los datos históricos.

Otro algoritmo con el que la regresión lineal estacional se puede comparar es el algoritmo regresión lineal múltiple que se está ejecutando con el valor ficticio de la pendiente y el mes de la variable independiente del año, pero también puede ejecutarse la regresión lineal estacional en los ciclos estacionales que son más cortos o largos que un año.

Si el componente estacional se utiliza o no depende del resultado del test de estacionalidad realizado con este algoritmo de proyección. Si el resultado es negativo, se utiliza la regresión lineal sencilla. Si el resultado es positivo, el algoritmo estima el componente estacional además de calcular la tendencia con la regresión lineal.

El algoritmo utiliza el método de mínimos cuadrados normales para el cálculo de α y β : la precisión de cada valor previsto se mide con sus cuadrados residuales (distancia vertical entre el punto de datos y la línea ajustada) con el objetivo de hacer que el total de estas desviaciones al cuadrado sea lo más pequeño posible.

Múltiple

La regresión lineal es un planteamiento para modelar la relación lineal entre una variable dependiente y una o más variables independientes, como el precio, la temperatura o el PIB. El

pronóstico depende de los valores futuros de estas variables independientes, que se conocen o se pueden calcular. Se debe tener en cuenta que no debe existir interdependencias (colinealidad) entre las diferentes variables. La proyección se calcula utilizando las funciones lineales, y los parámetros de modelo desconocidos se estiman a partir de los datos.

En la aplicación de demanda, la regresión lineal es útil cuando varias condiciones externas deben tenerse en cuenta durante el cálculo de proyección (por ejemplo, la temperatura media en ciertos períodos de tiempo). El sistema busca coeficientes con los que multiplicar las variables independientes y así obtener resultados que están lo más cerca posible a los valores históricos. De esta forma se pueden predecir los valores futuros de las variables independientes y calcular una proyección para el futuro que tenga en cuenta estas variables.

La ecuación que representa a este modelo es la siguiente:

$$Y_i = (\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_{ni} X_{ni}) + e_i$$

Donde:

β_0 : es la ordenada en el origen, el valor de la variable dependiente Y cuando todos los predictores son cero.

β_i : es el efecto promedio que tiene el incremento en una unidad de la variable predictora X_i sobre la variable dependiente Y, manteniéndose constantes el resto de las variables. Se conocen como coeficientes parciales de regresión.

e_i : es el residuo o error, la diferencia entre el valor observado y el estimado por el modelo.

5.7. Métodos predictivos con redes neuronales artificiales

El ámbito de las redes neuronales y su grupo mayor, el Deep Learning, es complejo y amplio. Durante los últimos años, el interés y la aplicación de este tipo de modelos han experimentado tal expansión que se ha convertido en una disciplina por sí misma. Si bien entender bien sus fundamentos requiere de una cantidad notable de tiempo y práctica, esto no significa que se necesiten adquirir todos ellos para empezar a sacarles partido; del mismo modo que no es necesario conocer el funcionamiento de todos los componentes de un "smartphone" (procesador, antena, circuitos...) para utilizarlo de forma productiva. En este trabajo de título se presenta una introducción, más intuitiva que rigurosa, sobre los modelos de redes neuronales y de cómo se crean a través de la herramientas y librerías que ofrece Python.

Las redes neuronales son modelos creados al ordenar operaciones matemáticas siguiendo una determinada estructura. La forma más común de representar la estructura de una red neuronal es mediante el uso de capas (layers), formadas a su vez por neuronas (unidades, units o neurons). Cada neurona, realiza una operación sencilla y está conectada a las neuronas de la capa anterior y de la capa siguiente mediante pesos, cuya función es regular la información que se propaga de una neurona a otra. En la siguiente figura se muestra una representación gráfica de estas:

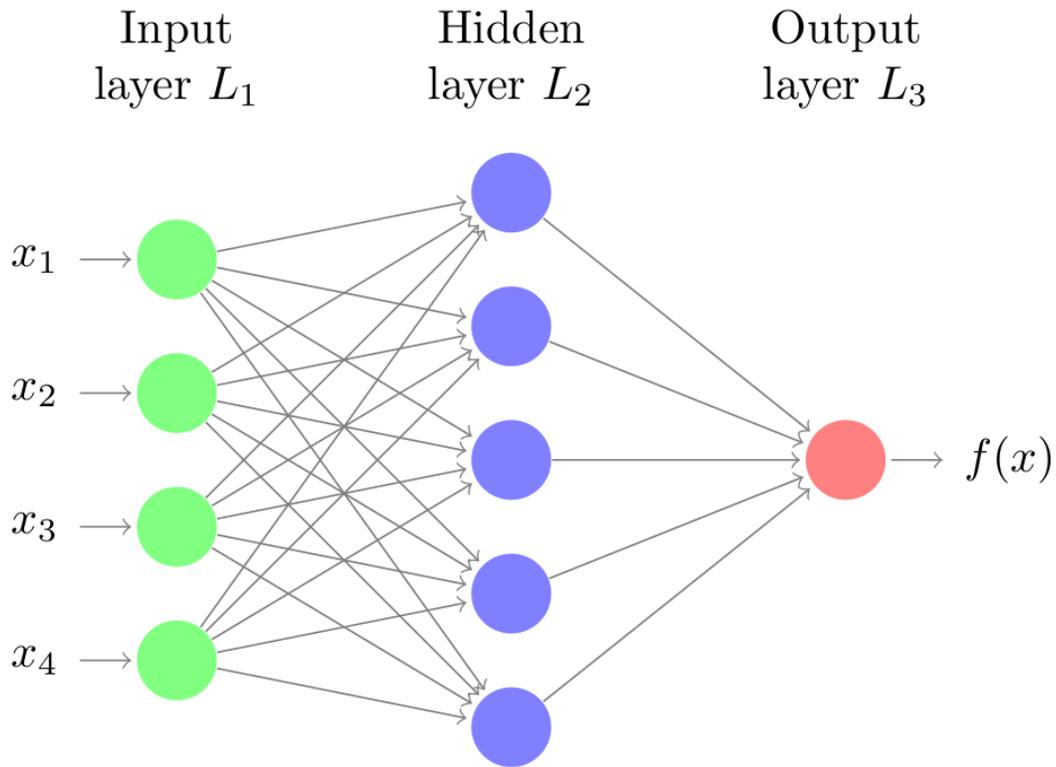


Figura 1. Representación de una red neuronal feed-forward (single-layer perceptron). Fuente: Computer Age Statistical Inference 2016.

La primera capa de la red neuronal (color verde) se conoce como capa de entrada o input layer y recibe los datos en bruto, es decir, el valor de los predictores. La capa intermedia (color azul), conocida como capa oculta o hidden layer, recibe los valores de la capa de entrada, ponderados por los pesos (flechas grises). La última capa, llamada output layer, combina los valores que salen de la capa intermedia para generar la predicción.

Para facilitar la comprensión de la estructura de las redes, es útil representar una red equivalente a un modelo de regresión lineal.

$$y = w_1x_1 + \dots + w_dx_d + b$$

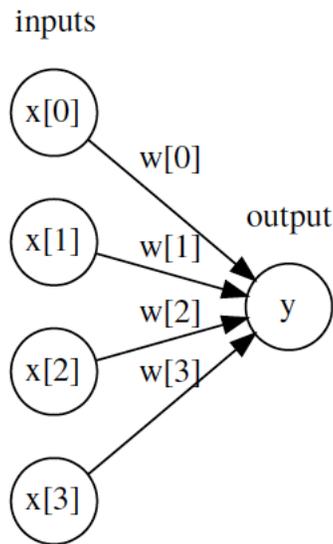


Figura 2. Representación de una red neuronal equivalente a un modelo lineal con 4 predictores. Fuente: COMS W4995 Applied Machine Learning

Cada neurona de la capa de entrada representa el valor de uno de los predictores. Las flechas representan los coeficientes de regresión, que en términos de redes se llaman pesos, y la neurona de salida representa el valor predicho. Para que esta representación equivalga a la ecuación de un modelo lineal, faltan dos cosas:

- El bias del modelo.
- Las operaciones de multiplicación y suma que combinan el valor de los predictores con los pesos del modelo.

La capa intermedia de una red tiene un valor de bias, pero suele omitirse en las representaciones gráficas. En cuanto a las operaciones matemáticas, es el elemento clave que ocurre dentro de las neuronas y conviene verlo con detalle.

La neurona

La neurona es la unidad funcional de los modelos de redes. Dentro de cada neurona, ocurren simplemente dos operaciones: la suma ponderada de sus entradas y la aplicación de una función de activación.

En la primera parte, se multiplica cada valor de entrada x_i por su peso asociado w_i y se suman junto con el bias. Este es el valor neto de entrada a la neurona. A continuación, este valor se pasa por una función, conocida como función de activación, que transforma el valor neto de entrada en un valor de salida.

Si bien el valor que llega a la neurona, multiplicación de los pesos por las entradas, siempre es una combinación lineal, gracias a la función de activación, se pueden generar salidas muy diversas. Es en la función de activación donde reside el potencial de los modelos de redes para aprender relaciones no lineales.

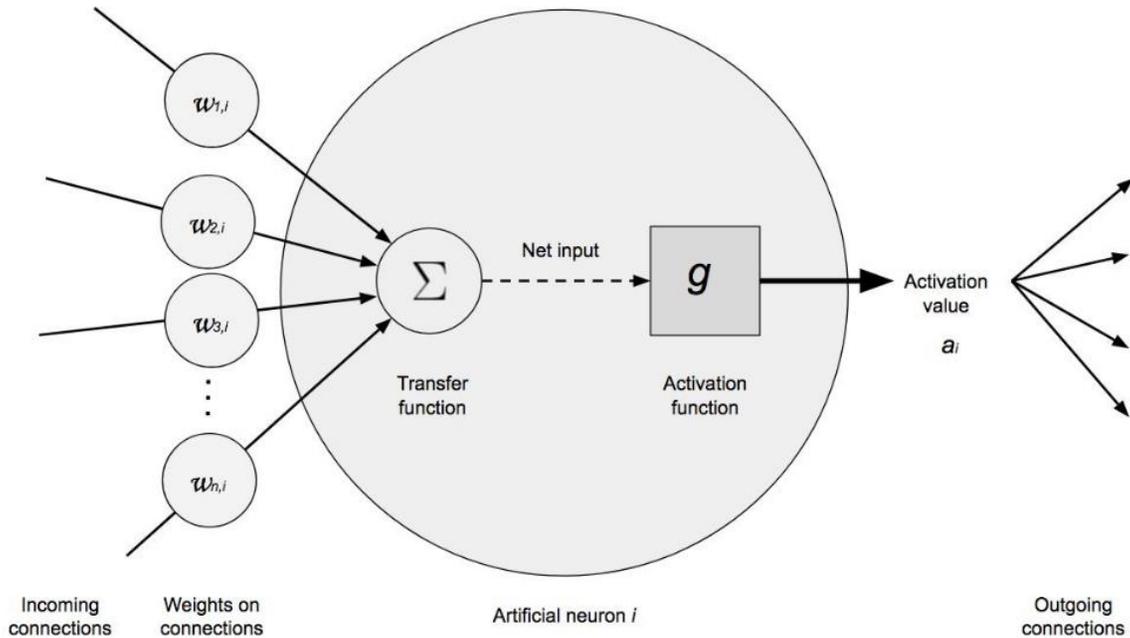


Figura 3. Representación de una neurona. Fuente: Deep Learning A Practitioner's Approach by Josh Patterson and Adam Gibson

La anterior ha sido una explicación intuitiva del funcionamiento de una neurona. Véase ahora una definición más matemática.

El valor neto de entrada a una neurona es la suma de los valores que le llegan, ponderados por el peso de las conexiones, más el bias.

$$\text{entrada} = \sum_{i=1}^n x_i w_i + b$$

En lugar de utilizando el sumatorio, esta operación suele representarse como el producto matricial, donde X representa el vector de los valores de entrada y W el vector de pesos.

$$\text{entrada} = XW + b$$

A este valor se le aplica una función de activación (g) que lo transforma en lo que se conoce como valor de activación (a), que es lo que finalmente sale de la neurona.

$$a = g(\text{entrada}) = g(XW + b)$$

Para la capa de entrada, donde únicamente se quiere incorporar el valor de los predictores, la función de activación es la unidad, es decir, sale lo mismo que entra. En la capa de salida, la función de activación utilizada suele ser la identidad para problemas de regresión y soft max para clasificación.

La neurona artificial se asocia muchas veces a una neurona del cerebro humano, pero ese no es su objetivo, sino el de crear algoritmos que puedan modelar distintos problemas. El poder de las redes neuronales proviene de su capacidad para aprender la representación de los datos de entrenamiento y de cómo los relaciona a la salida de la variable a predecir. Por este motivo las redes neuronales aprenden a hacer un mapeo matemáticamente. Esta estructura puede aprender a representar variables en diferentes escalas y combinarlas en variables más complejas.

Función de activación

Las funciones de activación controlan en gran medida que información se propaga desde una capa a la siguiente (forward propagation). Estas funciones convierten el valor neto de entrada a la neuronal, combinación de los inputs, pesos y bias, en un nuevo valor. Es gracias a combinar funciones de activación no lineales con múltiples capas (ver más adelante), que los modelos de redes son capaces de aprender relaciones no lineales.

La gran mayoría de funciones de activación convierten el valor de entrada neto de la neurona en un valor dentro del rango (0, 1) o (-1, 1). Cuando el valor de activación de una neurona (salida de su función de activación) es cero, se dice que la neurona está inactiva, ya que no pasa ningún tipo de información a las siguientes neuronas. A continuación, se describen las funciones de activación más empleadas.

- ReLu

La función de activación ReLu aplica una transformación no lineal muy simple, activa la neurona solo si el input está por encima de cero. Mientras el valor de entrada está por debajo de cero, el valor de salida es cero, pero cuando es superior de cero, el valor de salida aumenta de forma lineal con el de entrada.

$$ReLU(x) = \max(x, 0)$$

De esta forma, la función de activación retiene únicamente los valores positivos y descarta los negativos dándoles una activación de cero.

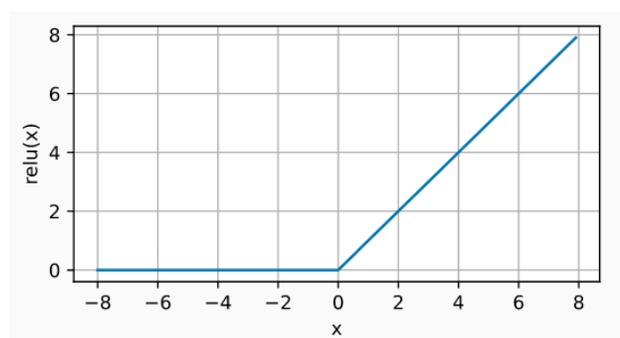


Figura 4. Representación función de activación ReLu.

ReLU es con diferencia la función de activación más empleada, por sus buenos resultados en aplicaciones diversas. La razón de esto reside en el comportamiento de su derivada (gradiente), que es cero o constante, evitando así un problema conocido como vanishing gradients que limita la capacidad de aprendizaje de los modelos de redes.

- Sigmoide

La función sigmoide transforma valores en el rango de $(-\infty, +\infty)$ a valores en el rango $(0, 1)$.

$$\mathit{sigmoid}(x) = \frac{1}{1 + \exp(-x)}$$

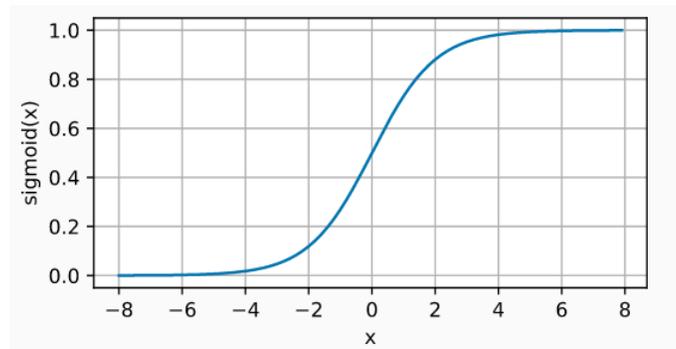


Figura 5. Representación función de activación Sigmoide.

Aunque la función de activación sigmoide se utilizó mucho en los inicios de los modelos de redes, en la actualidad, suele preferirse la función ReLU.

Un caso en el que la función de activación sigmoide sigue siendo la función utilizada por defecto es en las neuronas de la capa de salida de los modelos de clasificación binaria, ya que su salida puede interpretarse como probabilidades.

- Tangente hiperbólica

La función de activación Tanh, se comporta de forma similar a la función sigmoide, pero su salida está acotada en el rango $(-1, 1)$.

$$\mathit{tanh}(x) = \frac{1 - \exp(-2x)}{1 + \exp(-2x)}$$

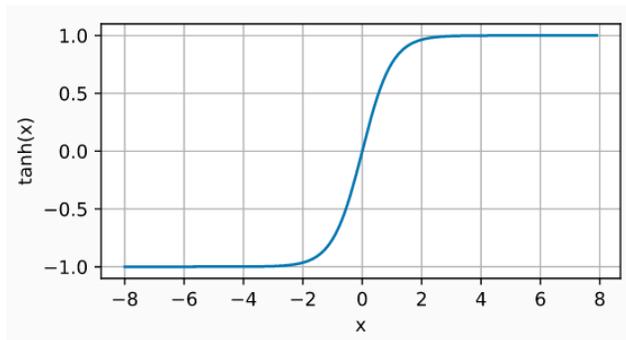


Figura 6. Representación función de activación tangente hiperbólica.

Cabe destacar que, sin las funciones de activación, las redes neuronales solo pueden aprender relaciones lineales.

Función de pérdida

Para poder controlar la salida de una red neuronal, primero es necesario poder medir la diferencia entre el valor esperado y el valor calculado por la red. Este es el trabajo de la función de pérdida. La función de pérdida toma las predicciones de la red y los valores verdaderos de la variable, luego calcula la diferencia entre sí, midiendo de esta forma el comportamiento de la red para este ejemplo específico. Este valor calculado es entonces utilizado como realimentación para ajustar los pesos, de forma que se reduzcan las diferencias.

Dicho ajuste es el trabajo del algoritmo de optimización. En el presente trabajo de título la función de pérdida que se utiliza es el MSE.

Algoritmo de Optimización

El algoritmo de optimización se utiliza para minimizar o maximizar una función objetivo. Determina la actualización de la red neuronal, minimizando la función de pérdida. Su implementación se basa en la técnica de optimización 'Descenso del Gradiente'. En el caso de las redes, la derivada parcial del error respecto a un parámetro (peso o bias) mide cuánta "responsabilidad" ha tenido ese parámetro en el error cometido. Gracias a esto, se puede identificar qué pesos de la red hay que modificar para mejorarla. El siguiente paso necesario, es determinar cuánto y cómo modificarlos (optimización).

Entrenamiento

El proceso de entrenamiento de una red neuronal consiste en ajustar el valor de los pesos y bias de tal forma que, las predicciones que se generen tengan el menor error posible. Gracias a esto, el modelo es capaz de identificar qué predictores tienen mayor influencia y de qué forma están relacionados entre ellos y con la variable respuesta.

Para el entrenamiento de una red neuronal mínimamente se requiere de las capas de la red, los datos de entrada, los objetivos de salida, la función de pérdida y el algoritmo de optimización. En este esquema, la red compuesta por capas conectadas mapea los datos de entrada a las predicciones obtenidas. Luego, la función de pérdida compara dichas predicciones con los objetivos

de salida esperados, obteniendo así una medida del error. El algoritmo de optimización utiliza dicha medida para actualizar los pesos sinápticos de la red neuronal. Inicialmente a los pesos de la red se le asignan valores aleatorios, pero a partir de cada muestra de datos que la red procesa la función de pérdida se minimiza, luego la diferencia entre los valores de salida esperados y calculados se minimiza, y los pesos sinápticos se ajustan. Es decir, que los pesos de la red se actualizan a partir de los errores calculados para cada muestra. Esto representa en términos generales el proceso de entrenamiento. El resultado final, luego de suficientes iteraciones es un modelo que representa o generaliza los datos involucrados.

La idea intuitiva de cómo entrenar una red neuronal es la siguiente:

1. Iniciar la red con valores aleatorios de los pesos y bias.
2. Para cada observación de entrenamiento (X,Y) , calcular el error que comete la red al hacer sus predicciones. Promediar los errores de todas las observaciones.
3. Identificar la responsabilidad que ha tenido cada peso y bias en el error de la predicción.
4. Modificar ligeramente los pesos y bias de la red (de forma proporcional a su responsabilidad en el error) en la dirección correcta para que se reduzca el error.
5. Repetir los pasos 2, 3, 4 y 5 hasta que la red sea suficientemente buena.

Si bien la idea parece sencilla, alcanzar una forma de implementarla ha requerido la combinación de múltiples métodos matemáticos, en concreto, el algoritmo de retropropagación (backpropagation) y la optimización por descenso de gradiente (gradient descent).

Datos de entrenamiento y testeo

Antes de iniciar el proceso de aprendizaje es necesario particionar los datos. Éstos se dividen en dos conjuntos: un conjunto para entrenamiento y un conjunto para testeo. El conjunto para entrenamiento es el más grande y contiene los datos que serán utilizados para la creación de los modelos de aprendizaje automático y el estadístico. Luego, el conjunto para testeo es utilizado para la evaluación o validación de los modelos creados.

6. Marco metodológico

En este capítulo, se presentan en detalle las metodologías estudiadas, seleccionadas y en parte aplicadas para dar cumplimiento al objetivo de este proyecto. Es importante destacar que la metodología final es de carácter híbrida que se crea a partir de un complemento entre dos enfoques de la ingeniería que pueden ser prácticos para resolver el problema planteado previamente, por una parte, se incorpora un enfoque de minería de datos, y por otra parte, un enfoque de rediseño de procesos de negocios.

Entre las metodologías conocidas tanto para el manejo de datos como para la extracción de información de estos, se estudió la metodología llamada CRISP-DM (Cross Industry Standard Process for Data Mining) la cual establece a través de un proceso jerárquico de trabajo consistente en un set de tareas en cuatro niveles de abstracción (de lo general a lo específico) [37], analizar datos y rediseñar procesos de negocios para cualquier tipo de industria en base al análisis.

Junto con lo anterior, y una vez identificado los principales factores que explican los errores en la estimación de demanda, se desarrolla un modelamiento de los procesos usando BPMN (Business Process Modeler Notation) identificando tanto las falencias como también las oportunidades de mejora en el proceso, se hace un estudio y modelamiento de los métodos predictivos utilizados y se elabora una propuesta de rediseño del proceso considerando el enfoque de Oscar Barros, la idea es proponer mejoras al proceso actual, y evaluar rendimientos de los distintos modelos predictivos en la planificación de la demanda de la CDS.

Se estudiaron diversos métodos desarrollados con el objetivo de encontrar información y patrones en los datos, para conseguir un uso inteligente y fácil. Entre estos:

1. SEMMA (Sample, Explore, Modify, Model, Asses)
2. DMAMC (Definir, Medir, Analizar, Mejorar, Controlar)
3. CRISP-DM (Cross Industry Standard Process for Data Mining)

Sin embargo, el de CRISP-DM ha tenido un mayor reconocimiento y uso en aplicaciones realizadas en proyectos de investigación como desarrollo empresarial como lo muestra la siguiente figura.

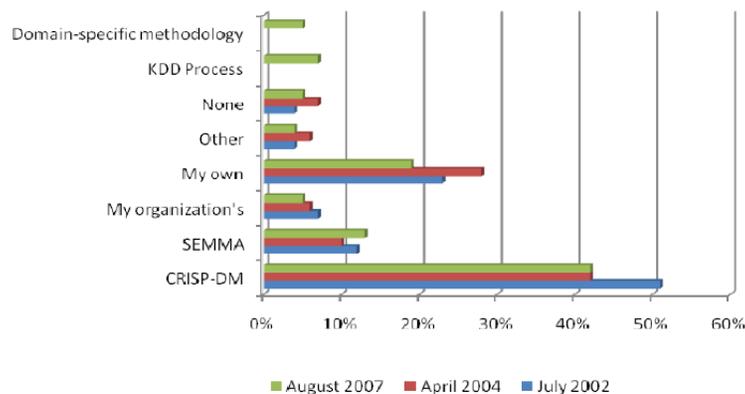


Figura 7. Encuesta de metodologías utilizadas en proyectos de Data Mining a nivel mundial, en año 2014. Fuente: [31]

6.1. Metodología CRISP-DM

La metodología CRISP-DM creada el año 1999 por el consorcio de empresas compuesto por NCR (Dinamarca), AG(Alemania), SPSS (Inglaterra), OHR (Holanda), Teradata y Daimler-Chrysler en la búsqueda de una guía de referencia sobre proyectos de Data Mining, proponen a partir de diferentes versiones de KDD, una metodología única con estructura jerárquica para el desarrollo de estos proyectos [39]

CRISP-DM se compone de cuatro niveles de abstracción ordenados desde el más general al nivel más específico.

- i. Fases
- ii. Tareas Generales
- iii. Tareas Especializadas
- iv. Instancias del Proceso

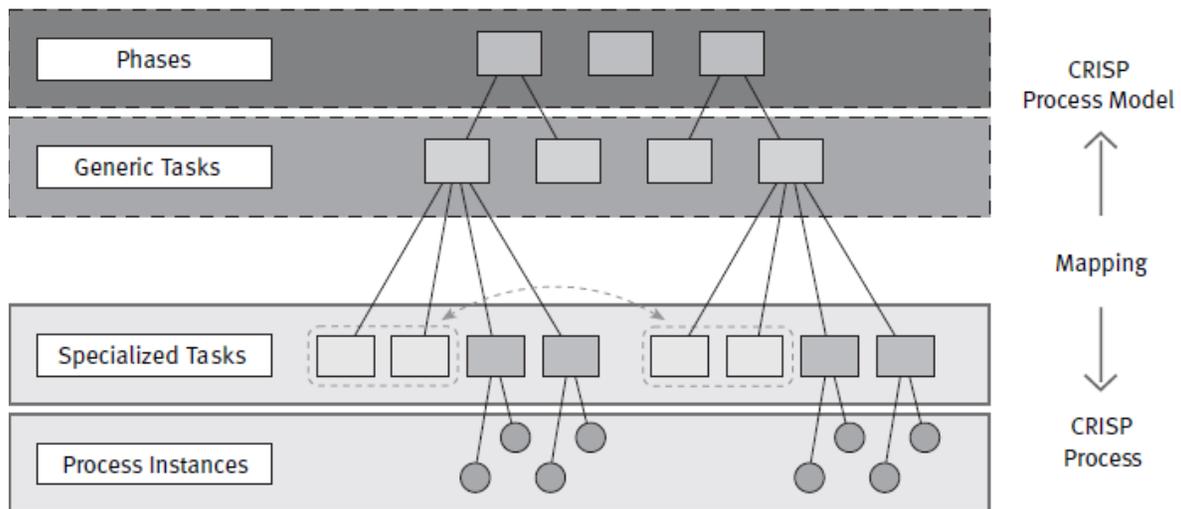


Figura 8. Niveles de abstracción del modelo CRISP-DM Fuente: [33]

Cada uno de estos niveles se compone de varias divisiones o tareas, las cuales son proyectadas en tareas específicas que finalmente se traducen en acciones desarrolladas para escenarios definidos. Esta estructura no necesariamente debe ser rígida, sino más bien las directrices expuestas en las tareas generales deben ser coherentes con la jerarquía de las tareas específicas y por supuesto con las acciones finales a ejecutar.

Desde el punto de vista genérico, el modelo CRISP-DM entrega una representación completa del ciclo de vida de un proyecto de minería de datos. Este se compone de seis etapas las cuales tienen tareas generales y específicas que se traducen finalmente en una implementación del modelo definido en diversas instancias del proceso.

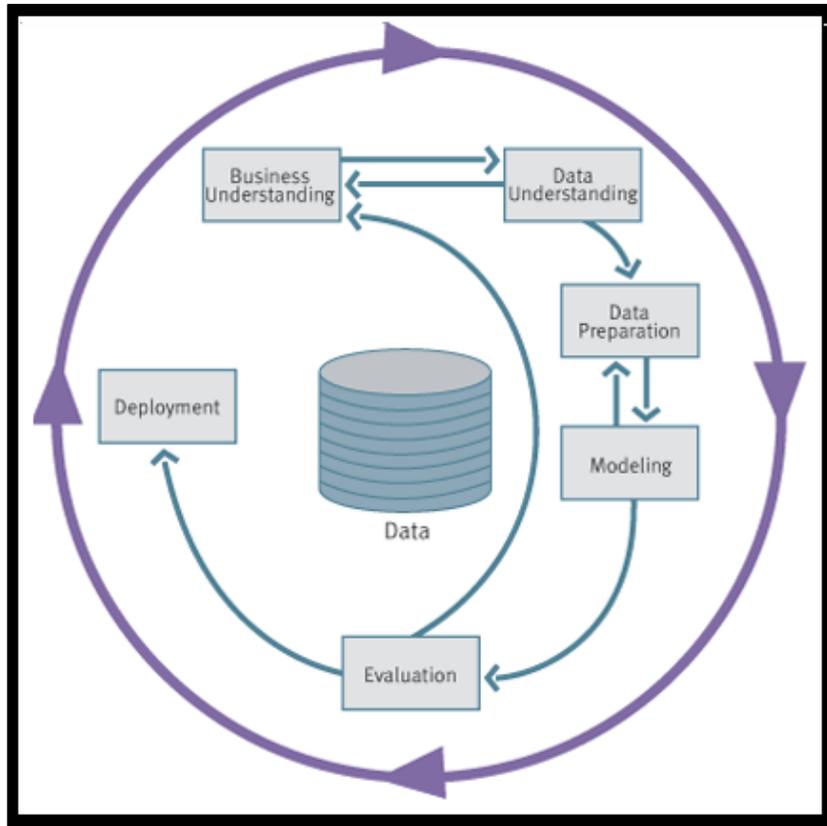


Figura 8. Modelo genérico CRIS-DM.

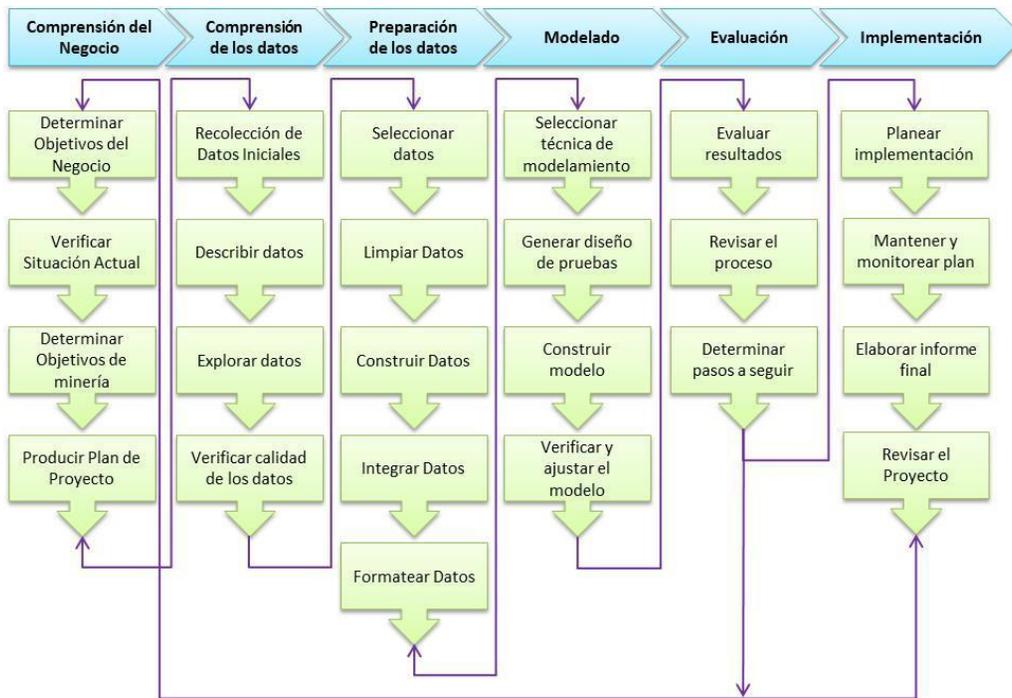


Figura 9. Fases y tareas del modelo CRIS-DM.

Fase 1: Comprensión del negocio

Esta fase consiste en la primera iteración en el proyecto a estudiar y analizar, es claramente la fase más importante del proceso ya que es la que aglutina la comprensión del problema, los objetivos del proyecto y la planificación de este. Si lo anterior presenta fallas o falta de comprensión del problema, cualquier algoritmo por sofisticado que sea, presentará deficiencias o fallas que requerirán modificaciones o mostrará resultados erróneos.

En esta fase, es muy importante poder materializar el problema identificado y en función del conocimiento del negocio, llevarlo a un problema de Data Mining que trace y permita alcanzar los objetivos dispuestos. Las principales tareas por definir en esta fase son:

- i. Determinar objetivos del negocio: consiste en definir el problema que se desea resolver, la necesidad de usar Data Mining y los criterios de éxito que tendrá el cumplimiento de estos objetivos.
- ii. Verificar situación actual: es la etapa o tarea que define como es el estado previo al uso de Data Mining o de la ejecución del proyecto en sí. En esta etapa de la fase, se debe tener claridad respecto de los siguientes ítems:
 - a. Conocimiento del problema
 - b. Datos requeridos para el análisis
 - c. Relación entre costo/beneficio de aplicar minería de datos para la solución de la problemática
- iii. Determinar los objetivos de minería: en esta tarea se materializan y concretan los objetivos del problema, tanto generales como específicos y se especifican los criterios de éxito o cumplimiento de estos objetivos.
- iv. Producir plan de proyecto: finalmente, se realiza un plan de proyecto que describa la metodología de desarrollo, las técnicas a emplear en cada una de las etapas y plazos.

Fase 2: Comprensión de los datos

Esta fase corresponde a entender la data disponible para enfrentar el problema, de forma de familiarizarse con este y elaborar las primeras directrices de procesamiento de los datos. Para esto, se requiere tanto definir los criterios de limpieza, como la recolección de datos faltantes u procesamiento requerido. Las principales tareas son:

- i. Recolección de los Datos Iniciales: Como primer paso de esta fase se debe levantar los datos y adecuarlos para el uso en las etapas siguientes, en paralelo a esto, se debe desarrollar un informe de recolección de datos el cual contenga la lista de datos recolectados, su fuente, como se realizó el levantamiento y los problemas detectados en el levantamiento con su respectiva solución.
- ii. Descripción de los Datos: Una vez obtenida la base de datos, ésta debe ser descrita en términos de volumen, campos que la componen, descripción de los campos y las características del formato, esto se realiza en un Reporte de Descripción de Datos.
- iii. Exploración de Datos: Posterior al reporte de descripción, se explora la base de datos en términos generales, se aplican pruebas estadísticas y tablas de frecuencias de modo de revelar

la estructura y las propiedades de los datos. Todas las pruebas generales realizadas a la base de datos se explican en un Reporte de Exploración de Datos.

- iv. Verificación de la Calidad de los Datos: En esta tarea se realizan pequeñas tareas simples sobre los campos de forma de chequear la consistencia de cada uno de estos, buscando valores fuera de rango o datos que puedan generar varianza sin significado para el problema. Esto se documenta en un Reporte de Calidad de Datos.

Fase 3: Preparación de los Datos

Ya con la base de datos constituida y revisada, se realiza la preparación de los datos para los análisis a realizar. Esta preparación considera desde la limpieza de la base de datos, extrayendo los datos que no se ocuparan en el análisis, hasta la creación de nuevas variables necesarias para el análisis y que no aparecen de forma explícita posterior al levantamiento realizado.

Esta fase tiene una relación directa con la fase siguiente (Modelamiento), esto ya que la extracción de datos, creación de variables y preprocesamiento de los datos dependen mucho del modelamiento que se está buscando realizar. En función de esto, las tareas que se realizan en esta etapa son:

- i. Seleccionar los Datos: De los datos obtenidos en la fase anterior, se selecciona un conjunto de ellos, los cuales serán los que constituirán el análisis que fundamentará el modelamiento, de forma tal que los análisis a realizar sean obtenidos solo de los datos necesarios, así se podrá obtener un mejor rendimiento y rapidez en los análisis.
- ii. Limpiar los Datos: luego de la selección de datos, y de forma complementaria a esta misma tarea de la fase anterior, se debe realizar una limpieza en base a los criterios que requiera esta nueva o sub-selección de los datos construidos en la fase anterior, de forma de asegurar que la calidad del análisis que sustentará al modelo en la fase siguiente.
- iii. Estructurar los Datos: posterior a la selección y limpieza de los datos a incluir en la modelación, se requiere dar una estructura en los casos donde se debe generar variables que no se tienen en la selección de los datos.
- iv. Integrar los Datos: dentro de la preparación de datos, también se puede reducir el número de variables a considerar dentro del modelamiento, esto puede ser mediante el uso de técnicas estadísticas, o por simple reducción de variables no necesarias.
- v. Formateo de Datos: en algunos casos, posterior a la estructuración de los datos se requiere un nuevo set de cambios, los cuales pueden ser simples (triming de datos, borrado de comas, etc.) a cambios más complicados (cambio en orden de los campos, cambio de ID's, tratamiento de valores fuera de rango, imputación de datos faltantes, etc).

Fase 4: Modelamiento

En esta fase, se debe realizar la elección y ejecución de la técnica de modelamiento seleccionada, para esto se debe tener en cuenta los criterios seleccionados en la fase de preparación de los datos,

el conocimiento del negocio para saber que técnica de modelamiento es la más adecuada, el manejo de la técnica de modelamiento a aplicar, entre otros.

Un antecedente importante para la selección de la técnica de modelamiento es la definición del plan de pruebas o escenarios que se someterá el modelo y la evaluación de este, generando indicadores que permitan escalar el grado de bondad que tendrá el modelo desarrollado. Los pasos son:

- i. **Seleccionar Técnica de Modelado:** Esta tarea consiste en seleccionar la técnica de modelamiento más adecuada al problema, para esto, se deben tener en cuenta las etapas anteriores y las tareas futuras, ya que el modelo debe ser el punto de convergencia donde la preparación de los datos es óptima para la técnica de modelamiento, y que la evaluación a obtener del modelo permitan resolver el problema de negocio en cuestión. Cabe señalar que la técnica de modelamiento no necesariamente es única, sino que puede ser un conjunto de técnicas que permita construir una estructura más realista del problema a resolver y entregue mayor cantidad de outputs para su evaluación.
- ii. **Generación Plan de Prueba:** Una vez generado el modelo, se debe contar con un plan de prueba que permita establecer la calidad y robustez de este. En algunos casos, es conveniente ejecutar el modelo sobre conjuntos de datos (distinto del set preparado en la fase anterior) de prueba para testear su validez en función del rango de error generado entre ambas pruebas, es decir, realizar ejecuciones del modelo con el conjunto de datos y compararlos con el set de datos de prueba para ver el rango de error arrojado por el margen de error entre ambas ejecuciones.
- iii. **Construcción del Modelo:** En la construcción del modelo, con la técnica de modelamiento escogida y los datos previamente preparados, se debe primero configurar los parámetros que usará el modelo para generar las predicciones, esto se realiza en forma iterativa sobre el conjunto de datos y se analiza la calibración de los parámetros comparando con resultados reales. En base a esto, es posible construir uno o más set de parámetros que se traducen en uno o más modelos, los cuales deben ser descritos y documentados en función de los parámetros escogidos.
- iv. **Evaluación del Modelo:** la evaluación del modelo consta de dos criterios de evaluación, el juicio experto del analista en Data Mining y el juicio ajustado al contexto del problema. Respecto del primero, es posible analizar los indicadores de éxito en la predicción realizada por el modelo en función de resultados esperados, respecto del segundo, existen resultados arrojados que pueden no ser coherentes con la realidad del problema o el negocio, los cuales sugiere normalmente un ajuste de parámetros o cambio de diseño en el modelo.

Fase 5: Evaluación

En esta fase se debe testear el desempeño y capacidad predictiva del modelo asociado a los datos, sin embargo, antes de dicha evaluación, se recomienda normalmente revisar el modelo con un enfoque del problema enfrentado, de modo de no olvidar alguna consideración que ocurra en la

realidad y no se esté considerando dentro del modelo y, por ende, no esté considerado en los resultados. Para efectos de evaluación, y a modo de complemento de análisis de resultados, es bueno clasificar los resultados del modelo en una matriz de confusión, la cual clasifica los resultados obtenidos en una matriz según resultado conocido, la diagonal de esta matriz arroja los resultados correctamente predichos por el modelo y el porcentaje de estos responde a una medida efectiva de desempeño del modelo. Las principales tareas por desarrollar en esta etapa son:

- i. Evaluación de resultados: Esta tarea consiste en revisar el modelo no en función de los resultados de las instancias ejecutadas, sino que, en términos de sus consideraciones, exactitud, robustez ante cambios de escenarios o set de datos, capacidad u horizonte de predicción, etc. De esta forma, se puede reportar un diagnóstico completo del modelo construido en función su capacidad y restricciones establecidas en la fase de Modelado en donde se establecen los supuestos para la construcción de este.
- ii. Revisión del Proceso: Posterior a la evaluación y documentación de esta, se realiza una revisión del proceso completo, desde la selección de datos hasta el desempeño del modelo, el cual tiene por objeto levantar posibles mejoras, ya que, en el desarrollo de este, existen a menudo criterios que pueden ser mejorados sin afectar en demasía los resultados obtenidos.
- iii. Determinar próximos pasos: Esta tarea se analiza en el caso de obtener resultados satisfactorios hasta esta tarea, en caso contrario, se revisa sobre qué fase conviene reestructurar el proyecto o si se debe iniciar desde cero para ampliar el campo de solución del problema seleccionado.

Fase 6: Implementación

En esta fase, ya con una visión cabal de los resultados obtenidos y de la estructura del modelo (criterios, supuestos y escenarios analizados), las recomendaciones o conclusiones surgidas de las predicciones hechas por el modelo constituyen un plan de implementaciones piloto en el negocio analizado, para esto, se desarrolla un plan de implementación el cual busca analizar primeramente, resultados en un sector acotado u parte del problema para comparar los resultados predichos y luego un plan de implementación más amplio que cierra el proyecto y la documentación final de este. Para esta fase, las principales tareas son:

- i. Plan de Implementación: La tarea de generar un plan de implementación de medidas como resultados del proyecto debe considerar entre ellas el impacto esperado de cada una. Es recomendable en ocasiones partir por implementaciones parcializadas de las medidas de modo de minimizar el efecto de encontrar problemas no levantados por los datos, ni por el modelo ni por el experto en el negocio. Este ranking de las medidas a implementar puede realizarse en función de costos o de lo invasivas que pueden ser estas medidas.
- ii. Plan de Monitoreo y Mantenimiento: Una vez definida las medidas a implementar, se debe monitorear y seguir estas medidas durante su fase de implementación, de forma de recibir feedback y retroalimentar las medidas propuestas y el modelo en general.

- iii. Informe Final: Ya con los resultados de la implementación de las medidas, se completa el informe final con estos resultados y las experiencias rescatadas del proyecto desarrollado, se documentan las conclusiones del proyecto y las posibles oportunidades de mejora aplicables al desarrollo realizado.
- iv. Revisión del Proyecto: En la presentación o revisión del informe final, se debe hacer un análisis somero de que puntos en el proyecto que tienen fueron bien desarrollados, que puntos fueron olvidados y que oportunidades de mejora son adaptables en el corto plazo dentro del proyecto, visto desde el punto de vista de un proyecto paralelo o sucesor al proyecto en revisión.

6.2. Metodología de rediseño de negocios de Oscar Barros

Por otro lado, se estudió la metodología de rediseño de negocios de Oscar Barros [35]-[36], que entrega un enfoque desde la Ingeniería de negocios, y entrega ideas para mejorar las estimaciones del proceso actual no sólo desde el punto de vista de generar un nuevo modelo a partir del enfoque CRIS-DM, sino que desde una perspectiva de procesos de negocio como lo es el S&OP de la CDS de Agrosuper que será explicado en los siguientes capítulos. Además que representa una metodología robusta para cualquier tipo de proyecto de mejoramiento de procesos de negocio.

La Ingeniería de Negocios será parte de la metodología base del presente trabajo, está desarrollada en el libro “Ingeniería de Negocios, Diseño Integrado de Servicios, sus Procesos y Apoyo TI” (Barros, 2015). Esta metodología provee las herramientas para el diseño formal de negocios relacionando la estrategia, el modelo de negocio, las capacidades implicadas, los procesos y las tecnologías de información de apoyo.

La siguiente figura muestra un modelo propuesto por la Ingeniería de Negocios para el diseño de negocios, basado en la estrategia y el modelo de negocio que una organización quiere poner en práctica.

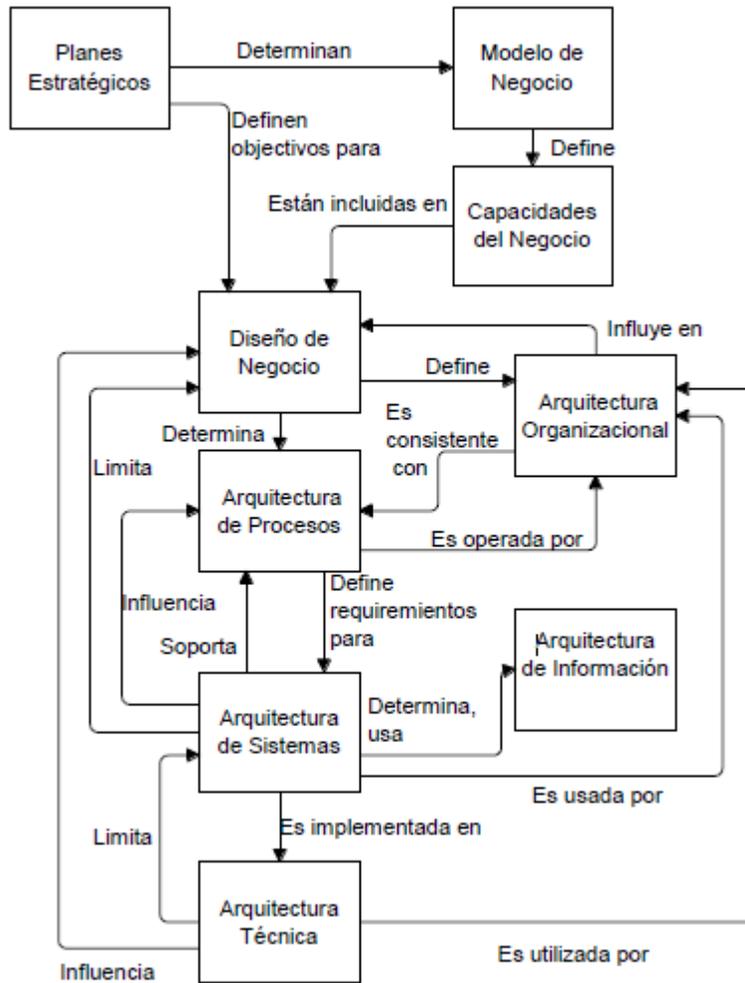


Figura 10. Ontología para el Diseño de Negocio

Teniendo en cuenta la Ontología para el Diseño de Negocio, se pueden abordar los problemas de diseño en los siguientes niveles:

Nivel 1: Diseño de Negocio que determina la estructura de componentes y sus relaciones, y la interacción con el medio ambiente que genera una Capacidad de Negocio¹, la cual provee un servicio valorado por los clientes de acuerdo con la estrategia y el modelo de negocio. Representa lo que debe hacer un negocio y no tiene un mapeo a las unidades organizacionales, área o producto.

Nivel 2: Diseño de la configuración y capacidad del negocio que incluye la determinación de los procesos que deben estar presentes para asegurar que el servicio definido en el punto anterior se proporciona en forma eficaz y eficiente.

Nivel 3: Diseño de procesos de gestión de recursos que incluye la gestión de personas, equipos y suministros necesarios para proporcionar la capacidad establecida en el punto anterior.

Nivel 4: Diseño de procesos de gestión de operaciones, incluye los procesos necesarios para la programación diaria de la demanda sobre los recursos para asegurar el nivel de servicio requerido y optimizar su uso.

Estos niveles pueden efectuarse de manera integral o parcial dependiendo de los objetivos que se persiguen. En el caso del presente trabajo, se aborda sólo el nivel 4 debido a que los niveles anteriores corresponden a la estructura de la organización que está fuera del alcance del proyecto. Sin embargo, aunque se hagan diseños parciales para algunos de los niveles definidos, el enfoque determina el uso de una metodología común que establece los determinantes de cualquier nivel de diseño.

Consta de las siguientes etapas que se explican a continuación:

Definir el Proyecto

Esta actividad pretende establecer con precisión cuáles son los procesos que deben ser rediseñados y los objetivos específicos que se tienen al enfrentar el cambio. Aquí, la idea fundamental es la de elegir y priorizar aquellos procesos que generen una mayor contribución al cumplimiento de los objetivos estratégicos de la organización. En esta etapa se establecen objetivos y se definen los ámbitos de procesos a rediseñar.

Entender Situación actual

En esta etapa se contempla:

Modelar la situación actual: Para esto se determinarán todas las actividades y agentes involucrados en el actual proceso de pronóstico de demanda. El proceso se representará a través del método de modelamiento BPMN, que utiliza un esquema eficiente para modelar sistemas muy complejos con muchas actividades y flujos, que también se utilizará para modelar el rediseño. La situación actual se obtendrá de la realización de las entrevistas, de la revisión bibliográfica existente de la empresa, entre otras actividades complementarias.

Validar y medir: etapa en la que se realiza una verificación de que los modelos representen fielmente lo que hoy día ocurre. Mediante entrevistas como las anteriormente señaladas se quiere lograr que las personas que operan y dirigen el proceso estén en acuerdo con que el modelo generado de la situación actual se ajusta a la realidad. Una vez producido un modelo aceptado, se procederá a medir los valores actuales de las variables que están involucradas en los objetivos del proceso. Esto requiere obtener información cuantitativa que lleve a darle valores numéricos a tales objetivos.

Rediseñar.

Establecer los cambios que deberían efectuarse en la situación actual y detallar cómo se ejecutarán los nuevos procesos. Se subdivide en:

Establecer direcciones de cambio: se entiende por dirección de cambio un conjunto de ideas globales que establecen la diferencia entre lo actualmente existente y el rediseño propuesto. Donde los patrones funcionan como algo normativo de cómo debieran ser.

Seleccionar tecnologías habilitantes: consiste en buscar y evaluar las tecnologías que hacen factible el cambio definido anteriormente. Para buscar y seleccionar tecnología lo más relevante es el tipo y el nivel de problema que se está enfrentando. Uno de los énfasis de la tecnología es facilitar el flujo de documentos y las decisiones asociadas al mismo.

Modelar y evaluar rediseño: consiste en realizar una representación de los nuevos procesos que implementarán el cambio establecido con anterioridad a través de un modelo BPMN. Para la evaluación económica se identificarán: los costos iniciales o de inversión, los costos recurrentes y los beneficios identificados al proyecto. Luego, se utilizarán indicadores como el VAN y la TIR para estimar rentabilidad de este.

Detallar y probar rediseño: en esta fase se deben detallar los procedimientos que deben ser realizados por personas como también los apoyos computacionales a través de un plan para su implementación. Para los procedimientos, el detalle tiene que ver con las prácticas de trabajo o rutinas que ejecutarán los colaboradores del proceso. En el caso de los apoyos computacionales, deben especificarse los elementos de hardware y software necesarios para implementar el proceso. En esta etapa se deberá crear un prototipo que demuestre la factibilidad del rediseño.

Implementar

En esta etapa se llevan a la práctica los procesos especificados anteriormente.

Construir software: Se deben definir las características y atributos específicos del software que apoyará el rediseño propuesto.

Implementar software: que significa poner en marcha definitiva la solución computacional diseñada. En esta etapa deben adquirirse todos los elementos de hardware y software faltantes; instalarlos en forma definitiva, para llegar a una operación rutinaria; y verificar la correcta operación de todos los elementos computacionales sobre los cuales correrá el software construido en la fase anterior.

Implementar procesos: lo cual conlleva el entrenamiento o capacitación de los participantes en el proceso, una marcha blanca para eliminar problemas de último minuto y una verificación de que el conjunto opera de acuerdo con lo diseñado y produce los resultados esperados.

6.3. Patrones de arquitectura y de procesos de negocios

Cualquier diseño hecho a partir de los Patrones de Negocio (PN) puede ser convertido en un correspondiente diseño de procesos por medio de la instanciación o especialización de uno de los Patrones de Arquitectura y de Procesos de Negocios. Éstos se basan en extensiva experiencia de diseño de procesos realizada en cientos de casos reales y comparten la idea de que existen cuatro agrupaciones de procesos llamados macroprocesos, que existen en cualquier organización, ellos son:

Macroproceso 1 (Macro 1): Conjunto de procesos que ejecutan la producción de bienes y/o servicios de la empresa, el cual va desde que se interactúa con el cliente para generar requerimientos hasta que éstos han sido satisfechos. Este macroproceso se asemeja a lo que otros autores denominan cadena de valor, sugerido por (Porter, *What is Strategy?*, 1996).

Macroproceso 2 (Macro 2): Conjunto de procesos que desarrollan las nuevas Capacidades que la empresa requiere para ser competitiva: los nuevos productos y servicios, incluyendo Modelos de Negocios, que una empresa requiere para mantenerse vigente en el mercado; la infraestructura necesaria para poder producir y operar los productos, incluyendo la infraestructura TI; y los nuevos Procesos de Negocios que aseguren la efectividad operacional y creación de valor para los clientes, estableciendo como consecuencia los sistemas basados en TI necesarios. En este macroproceso se lleva a cabo la innovación requerida.

Macroproceso 3 (Macro 3): Planificación del negocio, que comprende el conjunto de procesos necesarios para definir el curso futuro de la organización en la forma de estrategias, que se materializan en planes y programas.

Macroproceso 4 (Macro 4): Conjunto de procesos de apoyo que manejan los recursos necesarios para que los anteriores operen. Hay cuatro versiones que se pueden definir a priori: para recursos financieros, humanos, infraestructura y materiales.

Los Macroprocesos entregan una estructura integrada y coherente para el buen funcionamiento de la organización. Estos Macroprocesos descritos anteriormente se relacionan entre si mediante los flujos que se muestran en la figura, la cual se puede definir como una arquitectura genérica de macroprocesos que deben existir en cualquier empresa.

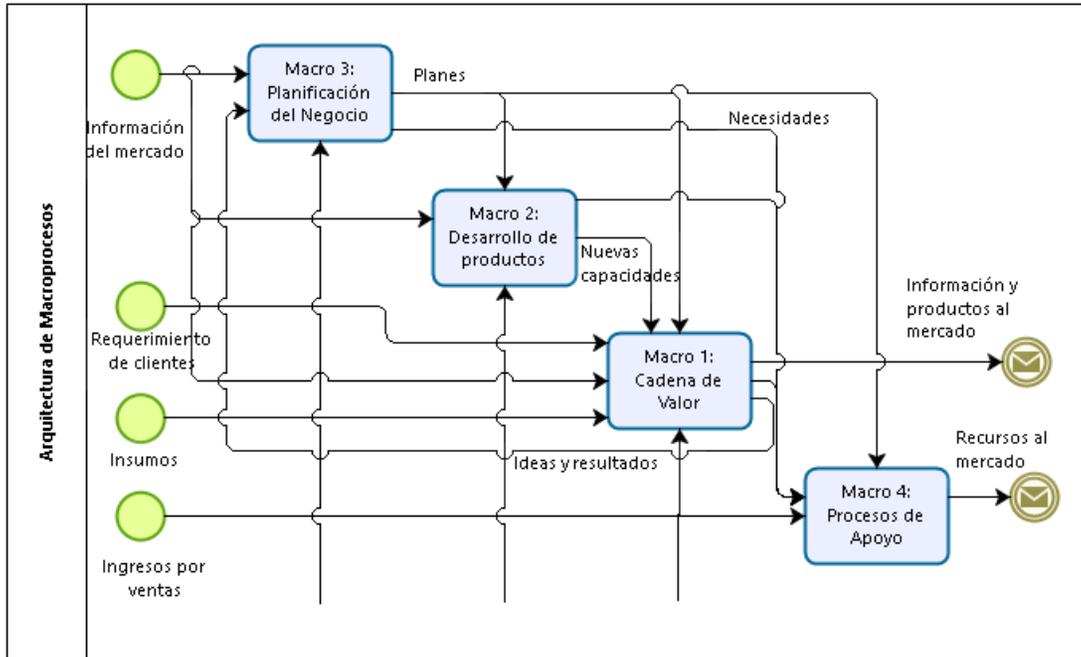


Figura 11. Arquitectura genérica de Macroprocesos

Además, es importante considerar las “Interrelaciones entre Macroprocesos”, las cuales se dan por medio de flujos, que pueden ser (1) Flujos materiales, es decir, insumos y otros recursos de proveedores, recursos desde el mercado, productos hacia el mercado, o recursos que fluyen internamente, o (2) Flujos de información, es decir, documentos, ya sea en papel o forma electrónica, que inducen la coordinación.

6.4. Patrones de Negocio

Los Patrones de Negocio (PN) son diseños generales de un negocio abstraídos de la experiencia, que describen como los elementos de un patrón pueden estructurarse en configuraciones de componentes que generan una Capacidad deseada. La necesidad de esa Capacidad se deriva de la Estrategia y del Modelo de Negocio y se relaciona con algún tipo de innovación que la organización desea realizar.

Se define un Patrón de proceso, como un modelo que señala cómo debería ser la estructura y funcionamiento de toda una clase de procesos que caen bajo un dominio en cuestión, basándose en una arquitectura, que sirve como guía para identificar componentes, relaciones y funciones, la cual debe ser complementada con el conocimiento del dominio al cual pertenece el proceso. La descripción detallada de los patrones de negocio están el libro “Ingeniería de Negocios: Diseño Integrado de Servicios, sus Procesos y Apoyo TI” (Barros, 2015). Se comentan a continuación.

PN1: Venta basada en conocimiento del cliente.

PN2: Creación de nuevas Value Streams (flujos de valor).

- PN3:** Aprendizaje interno para el mejoramiento de procesos.
- PN4:** Evaluación de desempeño para replanificación y mejora de procesos.
- PN5:** Innovación de productos.
- PN6** Uso óptimo de recursos.

El presente proyecto tiene un enfoque dado por el PN3 y PN4, debido a que son los más representativos para el contexto de este trabajo, con énfasis en el uso de Análítica para estudiar sistemáticamente el origen y posibles soluciones para los problemas del proceso utilizando datos de este, además del aprendizaje interno que se quiere lograr con la investigación y el análisis que se ha planteado y se planteará en los siguientes capítulos de este documento.

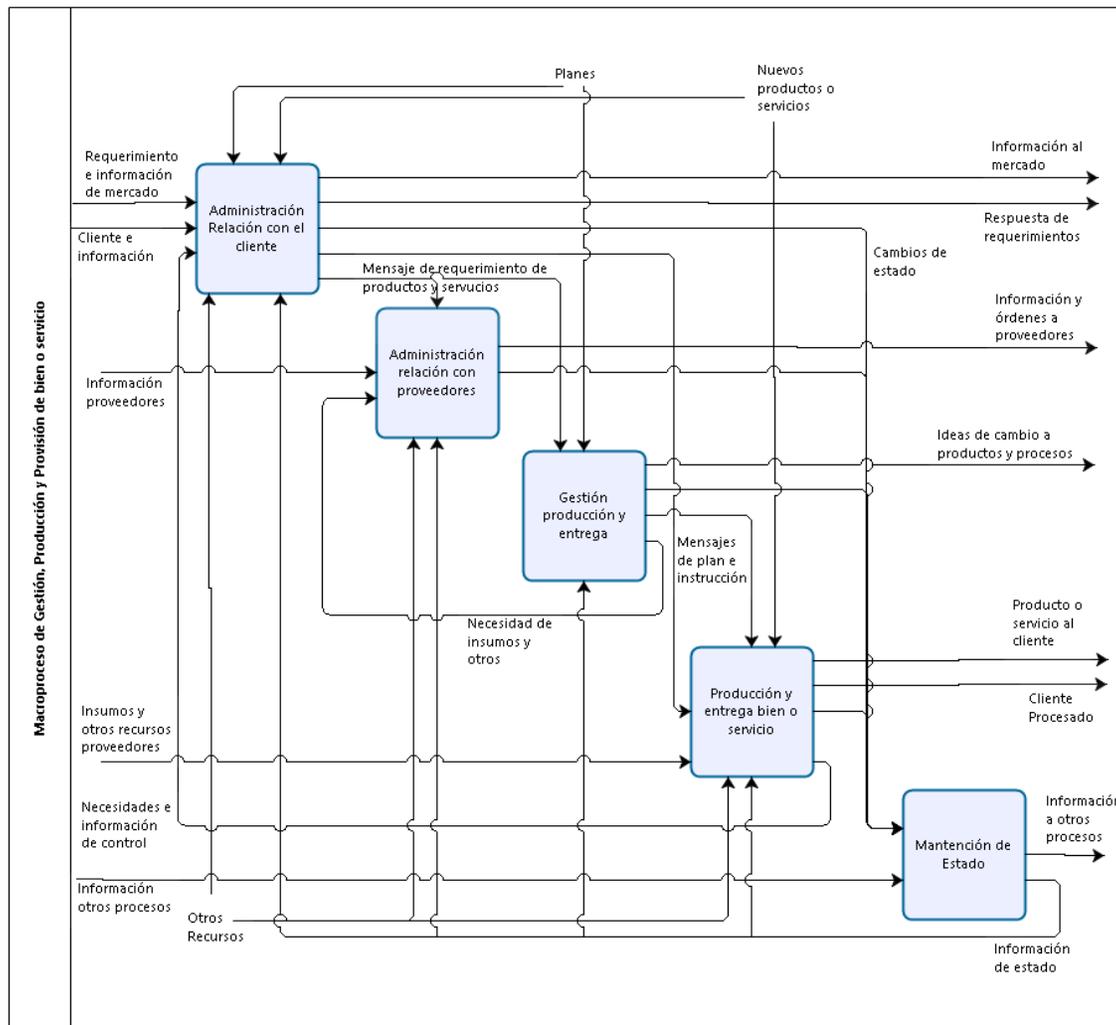


Figura 12. Patrón del Macroproceso Cadena de Valor

El primer patrón que se describirá es el del Macroproceso de Gestión, producción y provisión de un bien o servicio, el cual se muestra en la Figura 3, donde se observan sus flujos y actividades. Los procesos en los cuales se descompone este Macroproceso son los siguientes:

- Administración de la relación con el cliente: esta actividad se encarga de realizar las actividades de marketing y análisis de mercado, realiza la venta y atención a clientes, y decidir si es posible o no satisfacer los requerimientos de este. En la Figura 4 se muestra el patrón para este proceso.

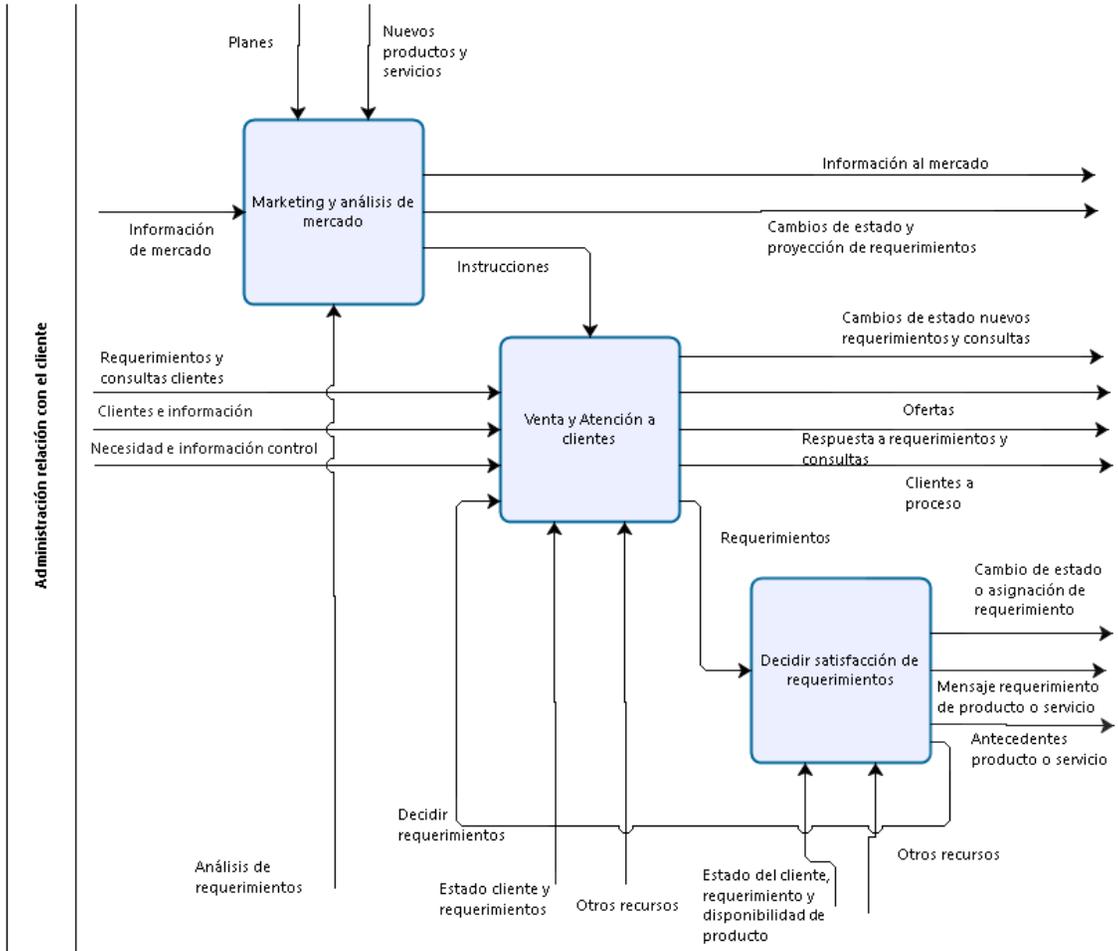


Figura 13. Patrón del proceso de Administración de la Relación con el cliente

- Administración de la relación con el proveedor: proceso que se encarga de interactuar con los proveedores de las materias primas constitutivas de los productos que ofrece la empresa.
- Gestión producción y entrega: este proceso, cuyo patrón se puede ver en la Figura 5, se encarga de generar los planes e instrucciones de producción y entrega, que le indica a la siguiente función qué, cómo y cuándo producir y entregar.
- Producción y entrega bien o servicio: satisface las necesidades de los clientes generando los productos que se entregan al mercado a partir de los insumos y recursos, cumpliendo con las instrucciones que entrega el proceso de "Gestión producción y entrega". En la Figura 6 se puede ver el patrón para este proceso.

- **Mantenimiento estado:** registra la situación o estado de todas las entidades que intervienen en el proceso, ya sea clientes, requerimientos de estos, proveedores, relaciones con éstos, recursos productivos, entre otros.

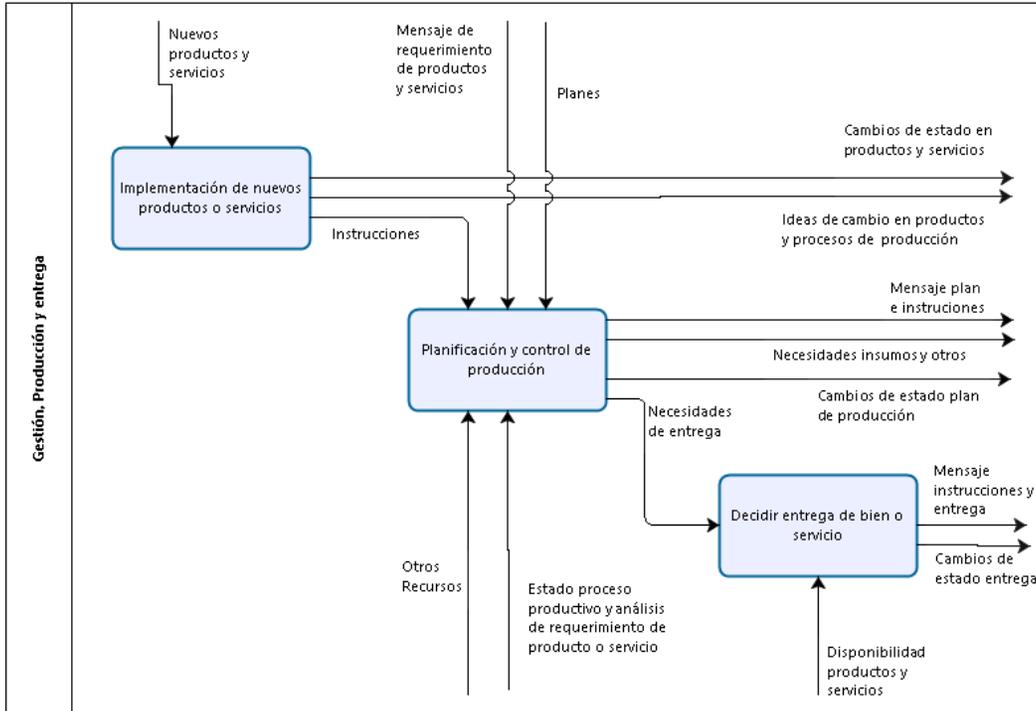


Figura 14. Patrón del proceso de Gestión, Producción y Entrega

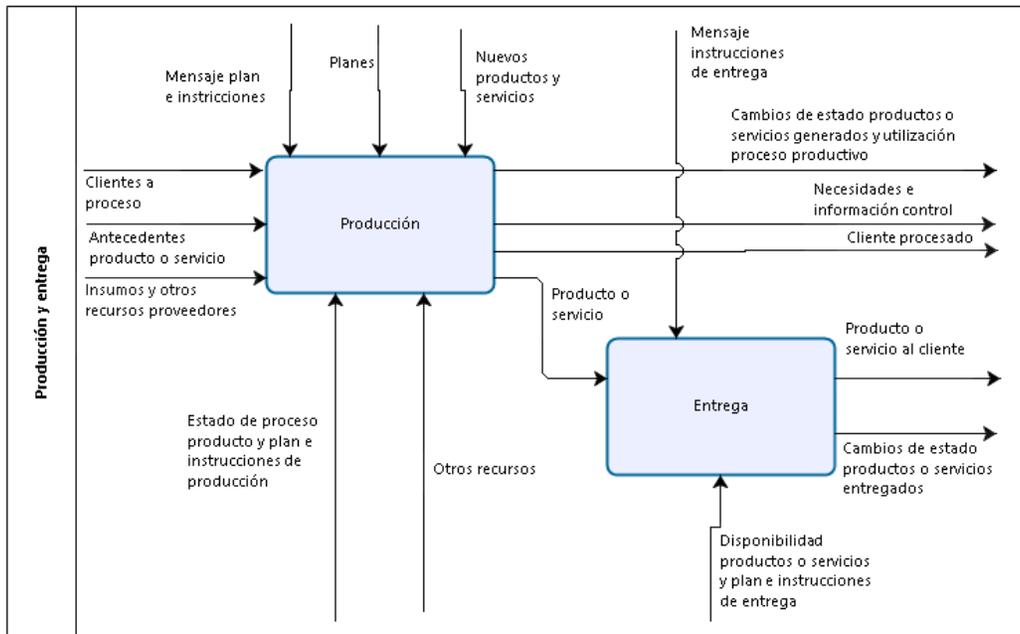


Figura 15. Patrón del proceso de Producción y Entrega.

6.5. Modelo y notación de procesos de negocio (BPMN)

El Object Management Group (OMG) ha desarrollado un estándar denominado Business Process Model and Notation (BPMN) que su principal objetivo es proporcionar una notación fácilmente comprensible por todos los usuarios de negocio, desde analistas que crean los borradores iniciales de los procesos, los desarrolladores responsables de implementar la tecnología que ejecutará esos procesos, por último, la gente de negocios que administrará y monitorizará esos procesos. De esta manera, BPMN crea un puente entre la brecha del diseño de los procesos de negocio y su implementación.

El modelado en BPMN se realiza mediante diagramas con un conjunto de elementos gráficos que representan actividades, subprocessos, compuertas de control de flujo, conectores, eventos, bandas, objetos de datos y mensajes entre otros. En la siguiente figura se pueden ver los elementos que conforman la notación BPMN, los cuales se pueden separar en las categorías que se describen a continuación:

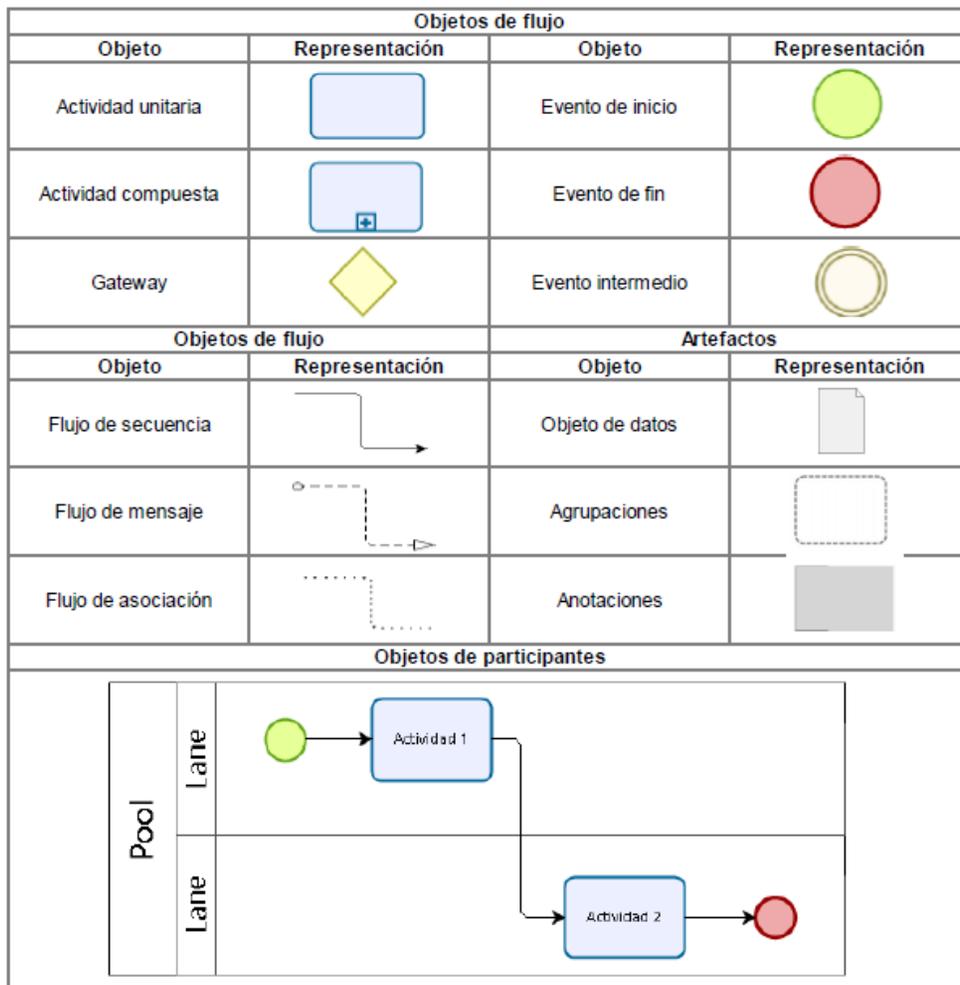


Figura 16. Elementos principales de BPMN

Objetos de flujo: son los elementos principales que definen los procesos. Los principales grupos de este tipo de elemento son:

- **Actividades:** se refiere a las tareas que se llevan a cabo dentro del proceso. Se representan a través de rectángulos.
- **Eventos:** son los hechos que suceden dentro del proceso, es decir, que no se llevan a cabo como las Actividades. Se representan a través de círculos.
- **Gateway:** se usan para modelar condiciones que se deben cumplir para realizar las Actividades o que sucedan Eventos. Se representan a través de rombos.

Objetos de conexión: se representan a través de flechas y sirven para modelar cómo es la relación y secuencia entre los objetos de flujo presentes en el proceso. Se subdividen en Flujo de Secuencia, Flujo de Mensaje y Flujo de Asociación.

Objetos de Participantes: se utilizan para agrupar los elementos antes descritos y representar las entidades coordinadas (Pool) y los responsables de los procesos (Lane). Estos elementos son rectángulos de mayor tamaño que contienen los elementos del proceso en su interior.

Artefactos: son distintos tipos de elementos que sirven para describir de mejor forma los procesos. Estos pueden ser Agrupación de elementos, Objetos de datos, Anotaciones, entre otros.

Cabe destacar que algunos de estos elementos cuentan con muchas variantes, en especial los objetos de flujo, lo que sirve para representar con mayor precisión distintas situaciones de procesos.

6.6. Variables de cambio

Se establece una serie de variables que sirven como un marco de referencia para determinar la dirección de cambio, la que se define como “un conjunto de ideas que establecen la diferencia entre lo actualmente existente y el rediseño propuesto” [3]. Estas variables de cambio permiten un diseño sistemático, pero no intentan detallar el rediseño, sino que más bien se centran en los conceptos más importantes para tener en cuenta. Las variables de cambio que se definen son:

- **Anticipación:** esta variable se refiere principalmente a establecer los mecanismos para anticipar los eventos futuros que influyan en la empresa. Por lo tanto, esta variable se relaciona principalmente con los procesos de planificación, es decir, establecer futuros requerimientos que deberá satisfacer el proceso y crear las condiciones para atenderlos satisfactoriamente.

- **Coordinación:** esta variable incluye prácticas como el uso de reglas, jerarquías, colaboración y partición, que permitan una correcta coordinación entre los miembros y procesos internos de la organización, así como también, con cliente y proveedores externos. Se debe tener especial consideración con los costos y beneficios que implica esta variable.

- **Prácticas de trabajo:** esta variable se refiere a la definición del rango de posibilidades, mecanismos, nivel de formalización o libertad en las reglas, procedimientos y rutinas para llevar a cabo una actividad, de forma tal que se cumpla el diseño de las otras variables del rediseño.
- **Integración de procesos conexos:** esta variable define el nivel de interacción entre los diferentes procesos que se van a rediseñar, y en particular entre los procesos que componen un macroproceso. Se refiere a integrar, por ejemplo, la obtención de recursos de los proveedores con su utilización en la generación de bienes y servicios, lo cual ofrece posibilidades de manejo optimizado de la cadena de abastecimiento.
- **Mantenimiento consolidada de estado:** esta variable se refiere al nivel de presencia de mecanismos que permitan obtener y proveer los datos que se necesitan para llevar a cabo las prácticas de trabajo, y comunicar los diferentes procesos y actividades, con el fin de facilitar la coordinación entre ellos.
- **Asignación de responsabilidades:** esta variable tiene relación con el grado de descentralización de las decisiones que se toman y las actividades que se realizan en los procesos. Se define también qué actividades serán realizadas por los miembros de la organización, o por personas externas a esta.
- **Apoyo computacional:** Es un resultado de las decisiones tomadas respecto a las variables anteriores, sobre todo, de las variables Prácticas de trabajo, Integración de proceso conexos y Mantenimiento consolidada de estado, mediante lo cual se definen los flujos computacionales y lógicas de negocio que se pueden automatizar.

6.7. Metodología final

Dadas las dos metodologías analizadas previamente, se propone una metodología híbrida que busca integrar una solución con ambos enfoques. Por una parte, la metodología de CRISP-DM para lograr desarrollar la fuente de innovación de este trabajo de título, el modelo de redes neuronales, pero no sólo eso, sino que todo lo relacionado al entendimiento, tratamiento, procesamiento, modelamiento y uso de los datos. Por otra parte, la metodología de rediseño de procesos de negocios para entregar una visión más orientada a la dinámica del proceso de S&OP, entendiendo que en el marco en el cual se realizan estas estimaciones es en un proceso más complejo, que requiere incorporar y tener considerado un enfoque de procesos.

Es importante también mencionar que este método fue diseñada y validado por el subgerente de planificación comercial (contraparte del proyecto), y por la líder de S&OP, a través de reuniones semanales en donde se presentaron las metodologías y se seleccionaron las actividades claves que requerían ser ejecutadas para el correcto desarrollo de este proyecto de mejora.

En la siguiente figura, se presentan las principales actividades ya realizadas y por realizar en los siguientes capítulos de este trabajo de título:

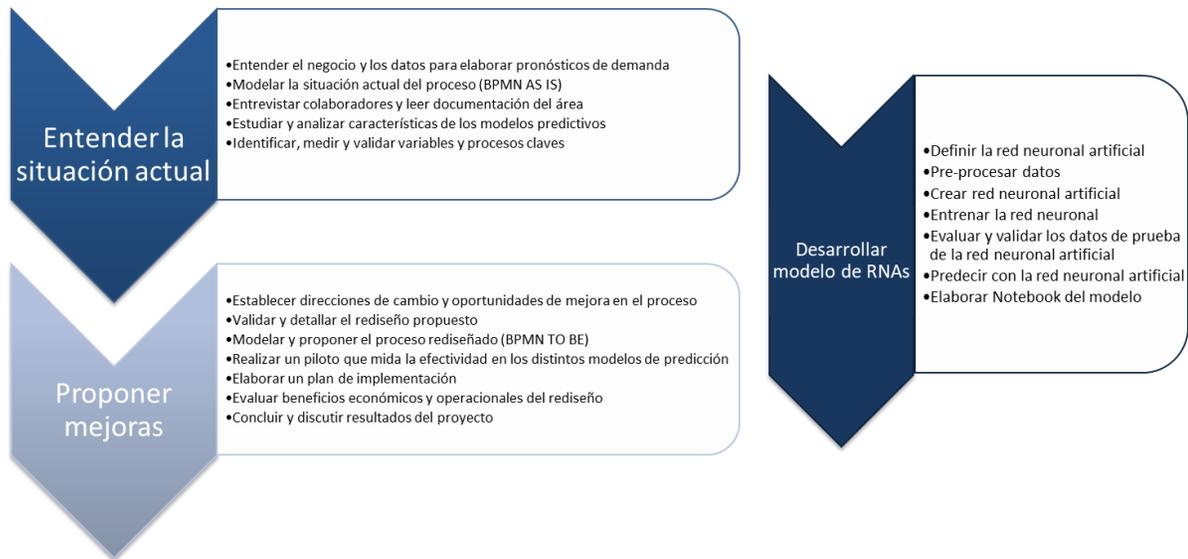


Figura 17. Metodología utilizada en este proyecto. Elaboración propia.

Además cabe destacar que las actividades no fueron desarrolladas en un orden cronológico entre actividades, sino más bien las actividades fueron realizándose de manera paralela en el tiempo, ajustándose también a los tiempos de colaboración que disponía el equipo y los sistemas de apoyo necesarios para ejecutar las tareas.

Dicho lo anterior, a continuación en el siguiente capítulo se presenta el desarrollo y modelamiento de la red neuronal artificial y adicionalmente de algunos de los modelos ofrecidos por SAP IBP (Auto-ARIMA/SARIMA) mediante los entornos provistos por Google Colaboratory, con lenguaje de programación y librerías ofrecidas por Python. Con lo anterior resuelto, se procede con la propuesta mejora del proceso y el desarrollo más integral de qué trata este proyecto.

7. Desarrollo del algoritmo de redes neuronales artificiales

A continuación, se presentará de manera más práctica lo que ha significado el desarrollo y el modelamiento de algunos de los métodos predictivos tradicionales, y en mayor medida, del modelo de RNA propuesto. Para esto se utilizará una serie temporal de ejemplo y se explicarán los resultados y la dinámica que hay detrás para predecir. En particular, se utilizará la serie temporal de demanda del SKU 1021560 el cual representa una categoría del sector cerdo (costillas). Cabe destacar que para cada una de las categorías se siguió la misma dinámica para modelar y extraer los resultados.

Se mostrarán también los resultados de las principales categorías de los sectores de cerdo y pollo, junto con las predicciones y MAPE que entrega el modelo de redes neuronales artificiales

7.1. Origen de datos

Se utilizará una data que contiene un conjunto de observaciones con información de demanda del SKU mencionado, medido en kilogramos. Específicamente los datos de prueba será una serie temporal con valores que comprenden el periodo entre la semana del 29-07-2018 hasta la semana del 25-07-2021, es decir, la frecuencia de los datos es semanal.

Dicho lo anterior, se presenta la siguiente tabla los datos donde se aprecia las observaciones tanto de la fecha como de la cantidad demandada respectivamente:

	Fecha	Demanda
0	2018-07-29	36480.44
1	2018-08-05	30646.96
2	2018-08-12	29335.32
3	2018-08-19	46824.59
4	2018-08-26	52309.91
...
152	2021-06-27	86497.45
153	2021-07-04	70875.53
154	2021-07-11	65174.86
155	2021-07-18	69803.69
156	2021-07-25	65111.30

Tabla 6. Datos o Serie de tiempo utilizada para modelar y hacer pronósticos de demanda

7.2. Software y librerías utilizadas

A continuación, se mencionan definiciones y detalles de los softwares y de las librerías utilizadas en este trabajo:

Python es un lenguaje de programación de código abierto orientado a objetos interpretado de alto nivel con semántica dinámica, estructuras de datos de alto nivel, combinadas con tipeo y enlace dinámicos. Es un lenguaje de scripting que puede ser utilizado para generar aplicaciones rápidamente. Otra de sus principales características es la facilidad de integración con librerías en otros lenguajes como 'C', lo cual lo convierten en un lenguaje de unión de componentes previamente existentes. Gracias a estas características Python tiene hoy una amplia comunidad siendo uno de los lenguajes más populares junto con otros como 'R' para el análisis de datos, computación interactiva y visualización de datos. Como desventaja, al ser Python un lenguaje de programación interpretado, su código correrá de forma más lenta que en un lenguaje compilado como C++.

NumPy es una librería dentro del ecosistema de Python utilizada para el cómputo numérico. Provee estructuras de datos y algoritmos que son utilizadas en la mayoría de las aplicaciones científicas gracias a su procesamiento rápido y eficiente.

Pandas es otra de las principales librerías dentro del ecosistema de Python que provee estructuras de datos de alto nivel y funciones para el análisis y estadística de información con la flexibilidad de lenguajes como SQL.

Matplotlib es una librería utilizada para la generación de visualizaciones de datos en dos dimensiones muy utilizada también en aplicaciones de ciencias de datos.

Statsmodels es una librería que contiene, entre otras cosas, algoritmos y funciones de visualización para el análisis económico y estadístico. En este trabajo se utilizan de esta librería el modelo estadístico VAR y los gráficos de autocorrelación.

Scipy es una librería compuesta de herramientas y algoritmos matemáticos. SciPy contiene módulos para optimización, álgebra lineal, integración, interpolación, funciones especiales, procesamiento de señales y otras tareas para la ciencia e ingeniería.

TensorFlow es una poderosa plataforma de código abierto desarrollada por Google para el cómputo numérico rápido y eficiente que contiene un amplio ecosistema de herramientas y librerías para aprendizaje automático y la creación y entrenamiento de redes neuronales. Tensorflow tiene soporte para su ejecución en paralelo a través de múltiples CPUs o GPUs y distribuida a través de distintos servidores. Esto permite el entrenamiento de redes de gran tamaño sobre grandes conjuntos de datos.

Keras es una interfaz de programación de aplicaciones (API, por su sigla en inglés) de alto nivel para la creación y el entrenamiento de modelos de aprendizaje automático que ofrece ventajas como una interfaz simple y la creación de modelos en forma modular. Estableciendo a Keras como una API de alto nivel para TensorFlow se logra simplificar el workflow y minimizar el desarrollo de modelos.

Google Colab es un entorno de trabajo interactivo y abierto que permite desarrollar código en Python de manera dinámica y completamente online, a la vez que integrar en un mismo documento tanto bloques de código como texto, gráficos o imágenes.

7.3. Modelamiento y desarrollo de algoritmos

Se presentan los algoritmos utilizados y programados en Python, en particular, de los modelos de Auto-ARIMA, Auto-SARIMA o Auto-ARIMA estacional y el modelo de redes neuronales artificiales.

Cabe destacar que antes de modelar y realizar las predicciones se hace un tratamiento para detectar valores atípicos con la prueba Cuantil de Tukey, esta se utiliza para encontrar los cuartiles máximo, mínimo, mediano, superior e inferior de los datos de acuerdo con las características de distribución de datos, y dibujar el cuadro.

$$IQR = Q_3 - Q_1$$

Donde, Q_3 es el tercer cuantil, Q_1 es el primer cuantil. Juzgue el valor atípico de acuerdo con la siguiente ecuación:

$$\begin{cases} x_n > Q_3 + 1.5 IQR \\ x_n < Q_1 - 1.5 IQR \end{cases}$$

De acuerdo con los datos el diagrama de caja cúbica se dibuja y se muestra en el gráfico 11. Los resultados muestran que hay seis valores atípicos en la secuencia de datos.

median=45628.33, q1=36411.90, q3=62958.77, iqr=26546.87, minimum=-3408.40, maximum=102779.07

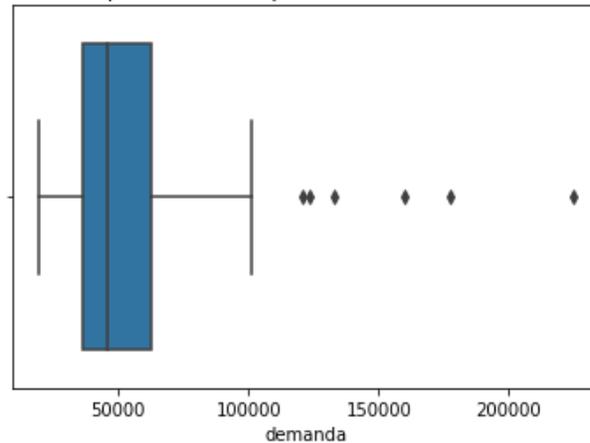


Gráfico 11. Gráfico de caja para determinar la cantidad de outliers de la serie temporal

Auto-ARIMA

Como se mencionó y explicó previamente en el marco teórico, el modelo de auto-ARIMA puede descomponer los dos primeros ordenes diferenciados de la serie temporal, esto se programó y se obtuvo como resultado los siguientes gráficos:

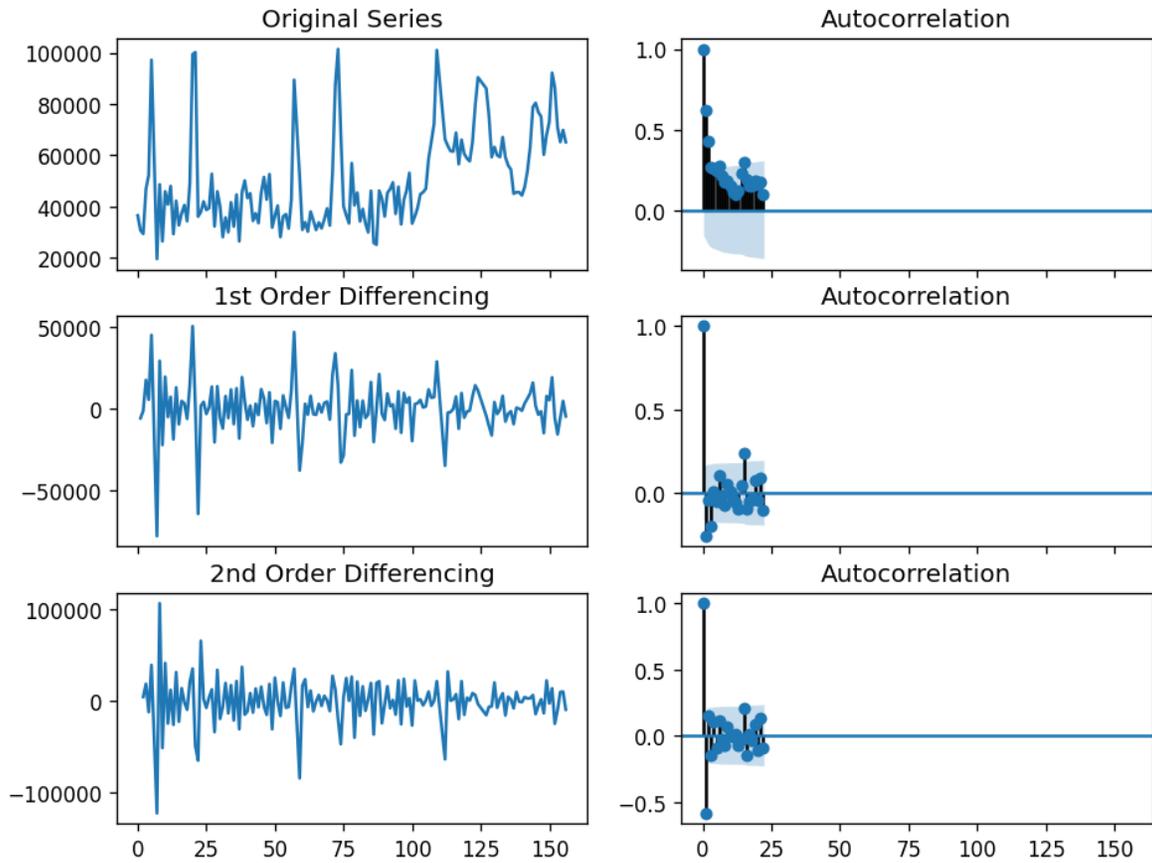


Gráfico 12. Gráfico representando la serie temporal y sus componentes diferenciadas.

Posterior a esto, se hace una prueba provista por librerías de Python para determinar el valor de p (cantidad de términos auto regresivos) y q (cantidad de promedios móviles) para el modelo de ARIMA, por consiguiente, se presentan los resultados obtenidos:

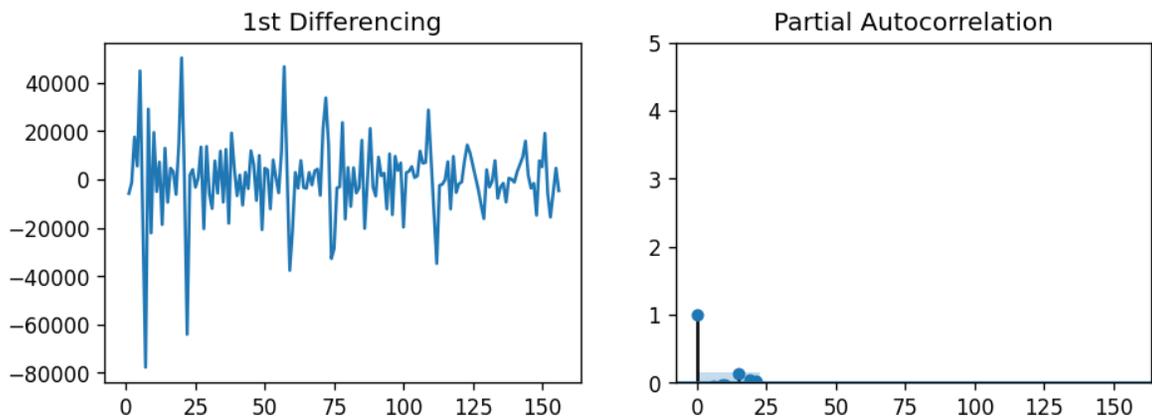


Gráfico 13. Gráfico representando el test de PACF para determinar el p de $AR(p)$.

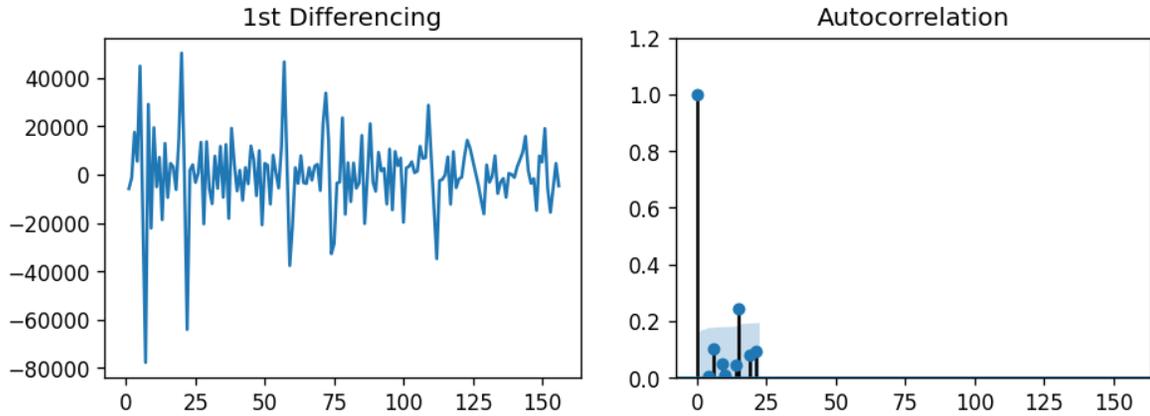


Gráfico 14. Gráfico representando el test para determinar el q de MA(q).

Luego, se procede a ejecutar el código para que el software haga el proceso de optimización del modelo de auto-ARIMA para encontrar los parámetros óptimos del modelo para predecir, esto se hace a través de la ejecución de distintas combinaciones de los parámetros p y q y el algoritmo elige el con mejor ajuste dado por el menor valor del indicador AIC el cual también fue explicitado en el apartado teórico de este proyecto, para este caso el modelo es el ARIMA (2,0,1):

```

Performing stepwise search to minimize aic
ARIMA(0,0,0)(0,0,0)[0] : AIC=3716.328, Time=0.07 sec
ARIMA(1,0,0)(0,0,0)[0] : AIC=3351.995, Time=0.02 sec
ARIMA(0,0,1)(0,0,0)[0] : AIC=3626.337, Time=0.03 sec
ARIMA(2,0,0)(0,0,0)[0] : AIC=3345.202, Time=0.04 sec
ARIMA(3,0,0)(0,0,0)[0] : AIC=3345.704, Time=0.07 sec
ARIMA(2,0,1)(0,0,0)[0] : AIC=3326.133, Time=0.13 sec
ARIMA(1,0,1)(0,0,0)[0] : AIC=3340.799, Time=0.05 sec
ARIMA(3,0,1)(0,0,0)[0] : AIC=3328.133, Time=0.20 sec
ARIMA(2,0,2)(0,0,0)[0] : AIC=3326.534, Time=0.20 sec
ARIMA(1,0,2)(0,0,0)[0] : AIC=3332.436, Time=0.11 sec
ARIMA(3,0,2)(0,0,0)[0] : AIC=inf, Time=0.47 sec
ARIMA(2,0,1)(0,0,0)[0] intercept : AIC=3326.186, Time=0.10 sec

Best model: ARIMA(2,0,1)(0,0,0)[0]
Total fit time: 1.527 seconds

SARIMAX Results
=====
Dep. Variable: y No. Observations: 151
Model: SARIMAX(2, 0, 1) Log Likelihood: -1659.067
Date: Fri, 10 Sep 2021 AIC: 3326.133
Time: 14:42:05 BIC: 3338.203
Sample: 0 HQIC: 3331.036
- 151
Covariance Type: opg
=====
coef std err z P>|z| [0.025 0.975]
-----
ar.L1 1.4831 0.081 18.342 0.000 1.325 1.642
ar.L2 -0.4834 0.081 -5.983 0.000 -0.642 -0.325
ma.L1 -0.9380 0.065 -14.539 0.000 -1.065 -0.812
sigma2 2.322e+08 9.04e-11 2.57e+18 0.000 2.32e+08 2.32e+08
=====
Ljung-Box (L1) (Q): 0.04 Jarque-Bera (JB): 100.82
Prob(Q): 0.85 Prob(JB): 0.00
Heteroskedasticity (H): 0.41 Skew: 0.65
Prob(H) (two-sided): 0.00 Kurtosis: 6.79
=====

```

Gráfico 15. Resumen de la optimización del proceso de auto-ARIMA.

Con el resultado anterior, ahora se procede a dividir la base de datos en un conjunto de entrenamiento y en otro de testeo o validación para poder evaluar la efectividad del modelo a la hora de predecir para posteriormente hacer proyecciones de valores futuros, se obtuvo los siguientes parámetros para el modelo:

ARMA Model Results

```

=====
Dep. Variable:          y      No. Observations:      120
Model:                 ARMA(2, 1)  Log Likelihood         -1322.766
Method:                css-mle    S.D. of innovations    14807.249
Date:                  Fri, 10 Sep 2021  AIC                    2655.532
Time:                  14:42:06      BIC                    2669.470
Sample:                0            HQIC                   2661.192
=====

```

	coef	std err	z	P> z	[0.025	0.975]
const	4.606e+04	2703.133	17.041	0.000	4.08e+04	5.14e+04
ar.L1.y	-0.2826	0.276	-1.025	0.305	-0.823	0.258
ar.L2.y	0.4217	0.134	3.154	0.002	0.160	0.684
ma.L1.y	0.7363	0.278	2.646	0.008	0.191	1.282

Roots

	Real	Imaginary	Modulus	Frequency
AR.1	-1.2409	+0.0000j	1.2409	0.5000
AR.2	1.9112	+0.0000j	1.9112	0.0000
MA.1	-1.3581	+0.0000j	1.3581	0.5000

Gráfico 16. Resultados del modelo de auto-ARIMA.

A continuación, se muestran los residuos y la densidad de la serie temporal y el proceso de optimización para lograr encontrar los parámetros óptimos en donde se aprecia la diferencia entre los valores predichos vs los reales:

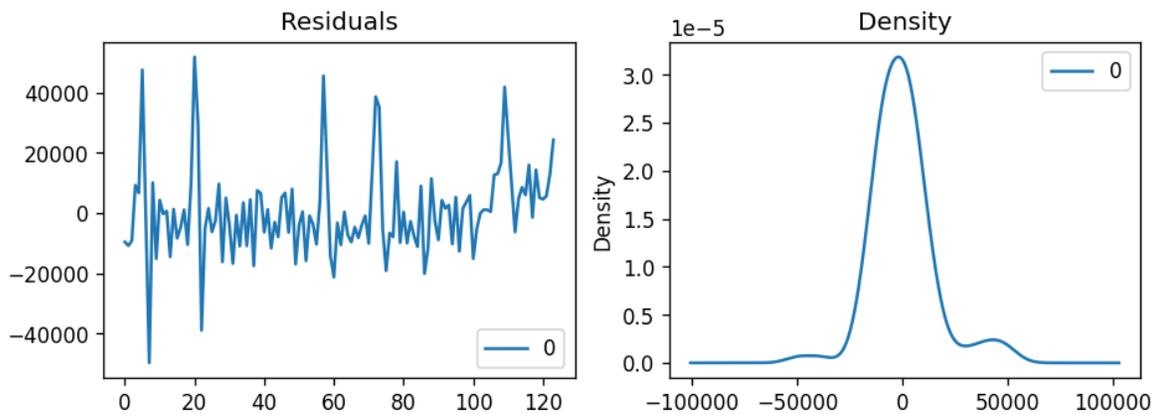


Gráfico 16. Residuos y Densidad de la optimización del proceso de auto-ARIMA.

Con todo lo anterior, ya se puede evidenciar la versión final de este modelo en el gráfico 17, en la cual se observa la serie original separada en un set de entrenamiento (color azul), un set de validación (color naranja) y las predicciones del modelo (color verde), en donde se calcularon métricas de precisión dando como resultado un MAPE de un 17,7% de precisión, lo que a simple vista no parece ser malo pues la meta establecida por la CDS de Agrosuper es de un 18%, por lo que se encuentra dentro del rango.

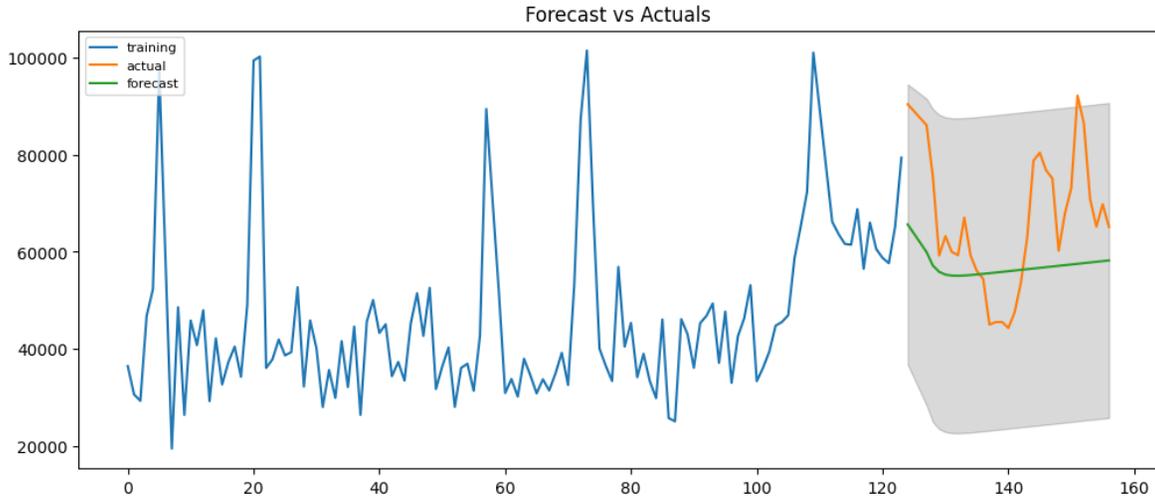


Gráfico 17. Forecast final vs real de la optimización del proceso de auto-ARIMA.

Auto-ARIMA estacional o Auto-SARIMA

A continuación, se presentan los resultados para obtener las predicciones con el método de Auto-SARIMA este método sigue la lógica anterior, pero su diferencia con el método de ARIMA, es que ahora se incorpora la estacionalidad dentro del análisis para hacer predicciones, por tanto ahora se presenta en el gráfico 18 la diferenciación original de la serie temporal y la diferenciación de la componente estacional de esta serie de tiempo:

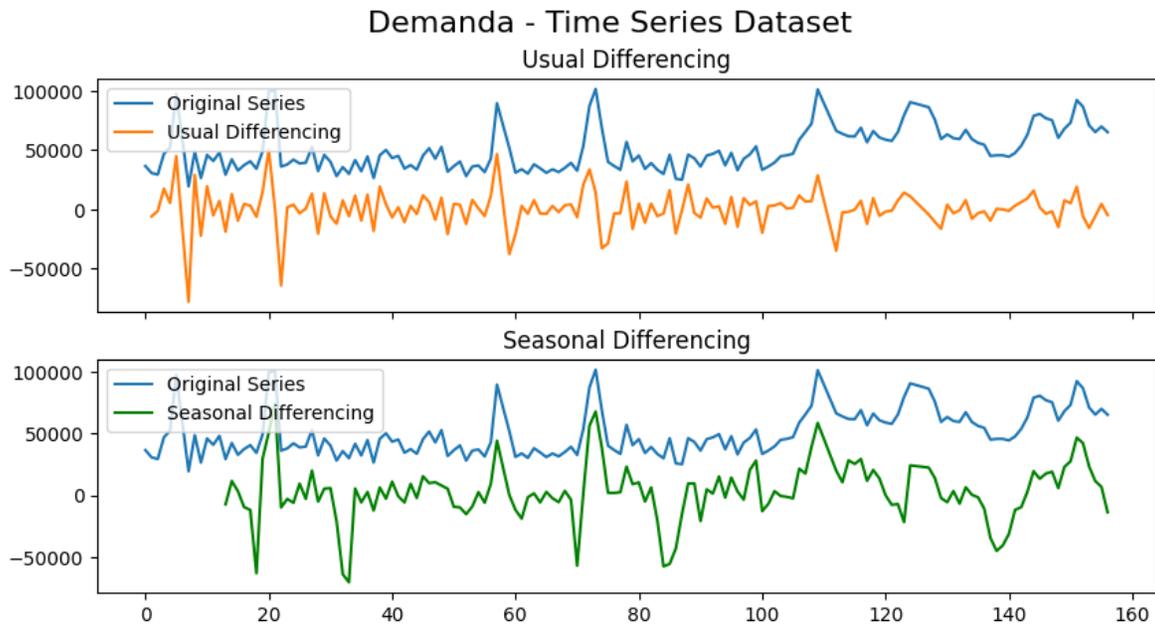


Gráfico 18. Diferenciaciones del modelo de auto-ARIMA estacional.

Luego, se realiza el mismo procedimiento anterior para obtener el modelo de Auto-SARIMA optimizado cuyos resultados se observan en el siguiente gráfico:

```

Performing stepwise search to minimize aic
ARIMA(1,0,1)(0,1,1)[52] intercept : AIC=inf, Time=4.51 sec
ARIMA(0,0,0)(0,1,0)[52] intercept : AIC=2245.155, Time=0.19 sec
ARIMA(1,0,0)(1,1,0)[52] intercept : AIC=2200.095, Time=3.34 sec
ARIMA(0,0,1)(0,1,1)[52] intercept : AIC=2208.925, Time=3.99 sec
ARIMA(0,0,0)(0,1,0)[52] intercept : AIC=2271.130, Time=0.28 sec
ARIMA(1,0,0)(0,1,0)[52] intercept : AIC=2226.952, Time=0.40 sec
ARIMA(1,0,0)(2,1,0)[52] intercept : AIC=inf, Time=16.08 sec
ARIMA(1,0,0)(1,1,1)[52] intercept : AIC=2192.935, Time=5.55 sec
ARIMA(1,0,0)(0,1,1)[52] intercept : AIC=inf, Time=3.75 sec
ARIMA(1,0,0)(2,1,1)[52] intercept : AIC=inf, Time=29.40 sec
ARIMA(1,0,0)(1,1,2)[52] intercept : AIC=inf, Time=22.85 sec
ARIMA(1,0,0)(0,1,2)[52] intercept : AIC=2190.457, Time=12.42 sec
ARIMA(0,0,0)(0,1,2)[52] intercept : AIC=2216.882, Time=13.35 sec
ARIMA(2,0,0)(0,1,2)[52] intercept : AIC=2193.156, Time=14.41 sec
ARIMA(1,0,1)(0,1,2)[52] intercept : AIC=2192.981, Time=14.35 sec
ARIMA(0,0,1)(0,1,2)[52] intercept : AIC=2208.103, Time=13.23 sec
ARIMA(2,0,1)(0,1,2)[52] intercept : AIC=2190.741, Time=25.21 sec
ARIMA(1,0,0)(0,1,2)[52] intercept : AIC=2207.901, Time=13.74 sec

Best model: ARIMA(1,0,0)(0,1,2)[52] intercept
Total fit time: 197.256 seconds

SARIMAX Results
Dep. Variable: y No. Observations: 151
Model: SARIMAX(1, 0, 0)x(0, 1, [1, 2], 52) Log Likelihood -1090.228
Date: Fri, 10 Sep 2021 AIC 2190.457
Time: 14:45:26 BIC 2203.432
Sample: 0 HQIC 2195.708
- 151

Covariance Type: opg
coef std err z P>|z| [0.025 0.975]
intercept 6473.5352 3150.517 2.055 0.040 298.038 1.26e+04
ar.L1 0.5733 0.174 3.289 0.001 0.232 0.915
ma.S.L52 -0.7839 0.242 -3.233 0.001 -1.259 -0.309
ma.S.L104 0.2502 0.261 0.957 0.338 -0.282 0.782
sigma2 3.235e+08 0.010 3.14e+10 0.000 3.24e+08 3.24e+08
Ljung-Box (L1) (Q): 0.75 Jarque-Bera (JB): 11.44
Prob(Q): 0.39 Prob(JB): 0.00
Heteroskedasticity (H): 0.89 Skew: -0.07
Prob(H) (two-sided): 0.73 Kurtosis: 4.66

```

Gráfico 19. Resultados del proceso de optimización del modelo de auto-ARIMA estacional.

En la misma lógica anterior, a continuación se presenta en el gráfico 20 la versión final del modelo de AUTO-SARIMA, en donde se tienen los mismos componentes y colores (entrenamiento, testeo y pronósticos). Además, se calculó que el MAPE para esta serie temporal es de un 29,49% lo cual es malo, por lo tanto, se evidencia empíricamente que este SKU no sirve proyectarlo con este tipo de modelo.

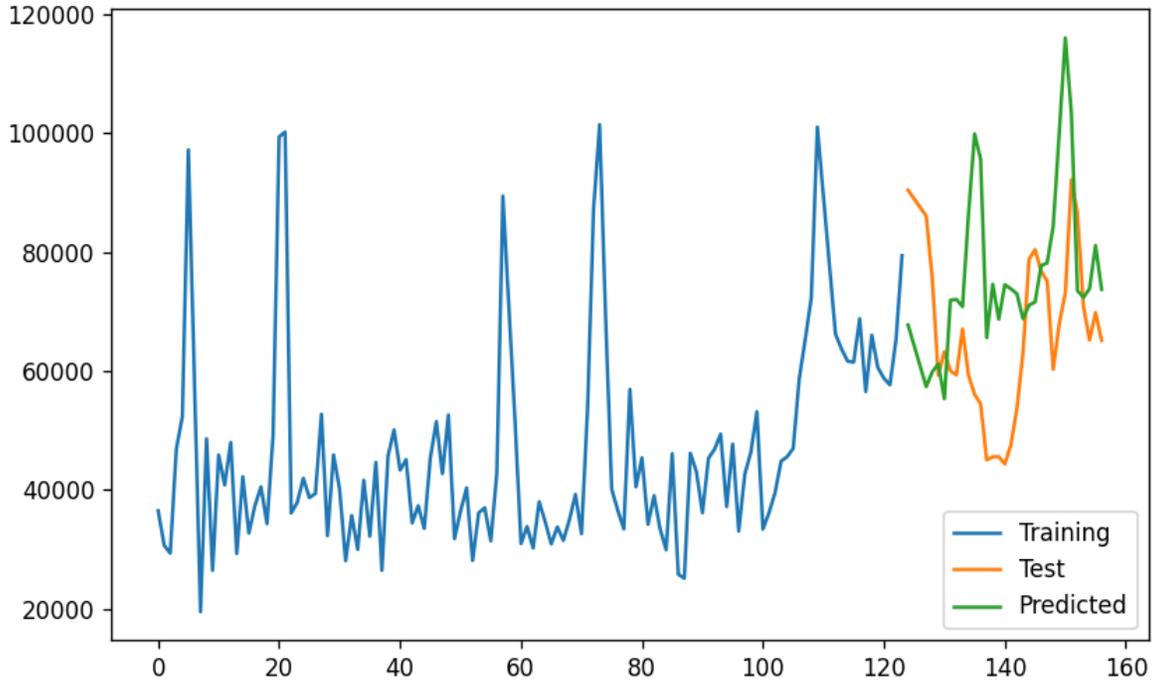


Gráfico 20. Resultado final del proceso de optimización del modelo de auto-ARIMA estacional.

Redes neuronales artificiales

El siguiente esquema resume los pasos para crear la red neuronal artificial y posteriormente validarla y predecir con ella:

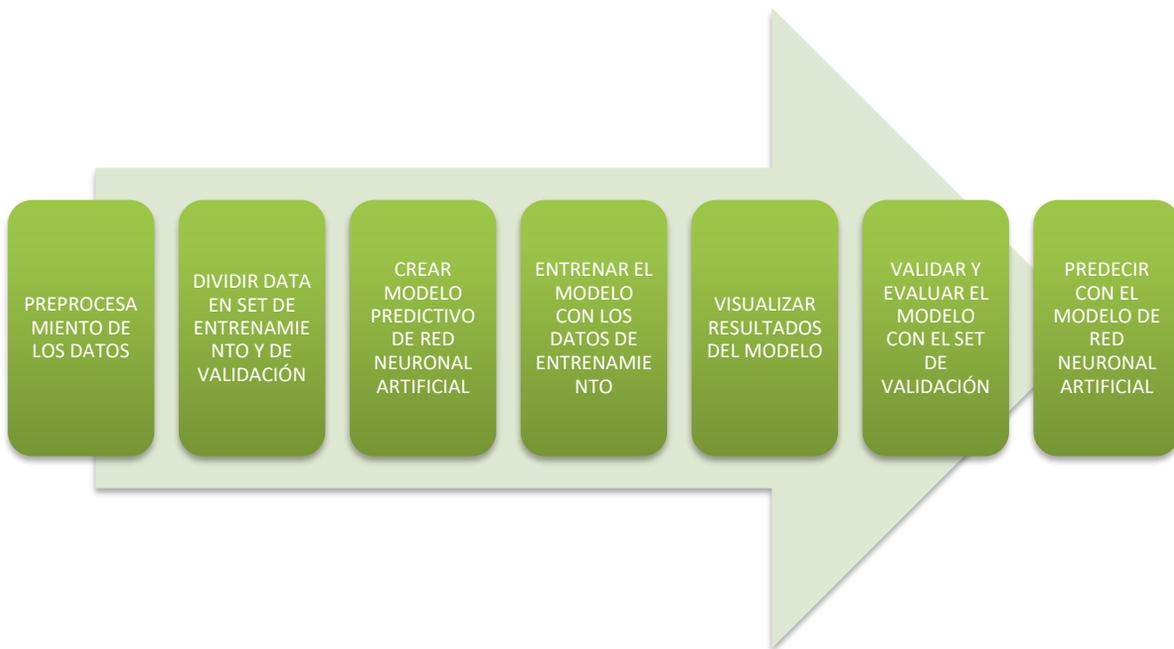


Figura 18. Pasos para crear la red neuronal artificial mediante el algoritmo en Python.

El siguiente gráfico se representa la serie original estudiada y ejemplificada:

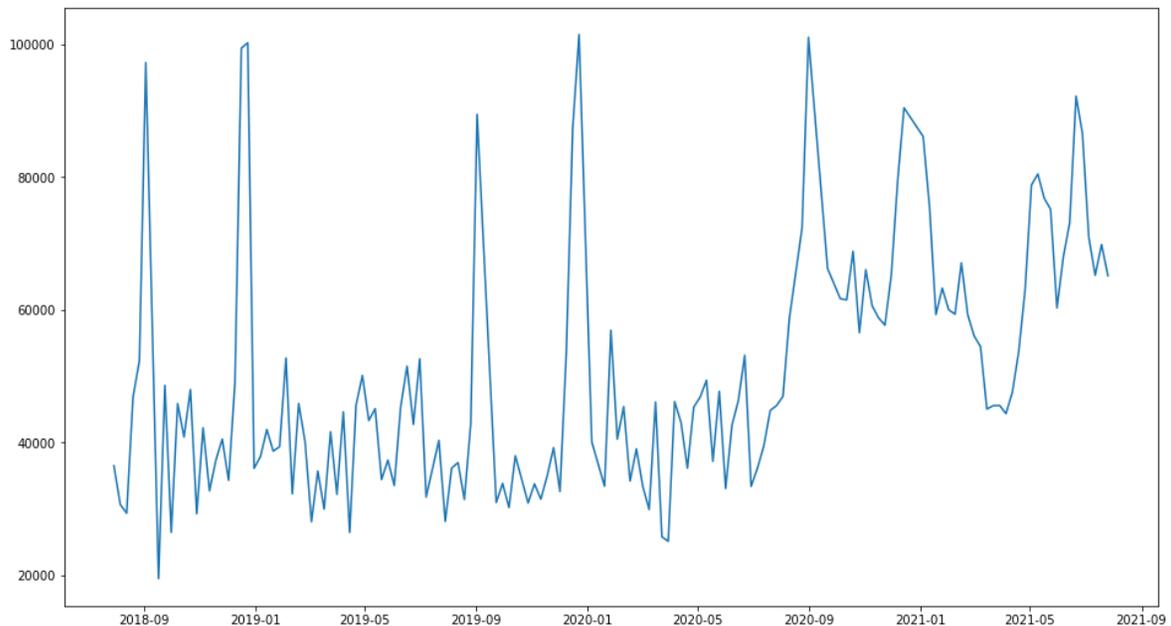


Gráfico 21. Serie temporal en estudio.

Por tanto, ahora se procede a mostrar y detallar el modelamiento de la red neuronal artificial mediante las librerías y entornos ofrecidos por Python y Google Colab, para esto se usará una arquitectura sencilla de red neuronal FeedForward (también conocido como MLP por sus siglas Multi-Layered Perceptron), no con un gran número de neuronas y como función de activación la llamada tangente hiperbólica presentada en el marco teórico, pues se entregan valores transformados entre -1 y 1.

Se hace una preparación de los datos, esto representa uno de los pasos más importantes de este tipo de ejercicios. En particular, se hace una alteración de los flujos de entrada del archivo importado, que contiene una columna con las unidades demandas, y esto se transformará en varias columnas. ¿Por qué hacer esto? Debido a que lo que se busca en realidad es tomar la serie temporal y convertirla en un problema de tipo supervisado, para poder alimentar la red neuronal y poder entrenarla posteriormente con backpropagation que es lo más habitual. Para esto, es necesario tener entradas y salidas definidas para hacer el entrenamiento.

Lo que se hace en este modelo y para el resto de las series temporales es tomar las últimas 10 semanas para obtener la onceava. Se podría entrenar con 3 o 5 semanas pero se optó para esta cantidad de semanas y por otra parte se podrían haber intentado predecir más de una semana futura, pero eso queda para otras investigaciones. Por tanto se tiene que las variables de entrada serán 10 columnas que representan las demandas en unidades de las últimas 10 semanas anteriores, y la variable de salida es el valor de la demanda de la onceava semana. Para esta

transformación se usa una función llamada `series_to_suervised()`, y antes de usar esta función se utiliza un `MinMaxScaler` para transformar el rango de valores entre -1 y 1 lo que a la red le favorece en términos para realizar cálculos. A continuación en la siguiente tabla se aprecia el set de datos de entrada

	var1(t-10)	var1(t-9)	var1(t-8)	var1(t-7)	var1(t-6)	var1(t-5)	var1(t-4)	var1(t-3)	var1(t-2)	var1(t-1)	var1(t)
10	-0.585376	-0.727779	-0.759798	-0.332863	-0.198959	0.896904	-1.000000	-0.289644	-0.830434	-0.356376	-0.479785
11	-0.727779	-0.759798	-0.332863	-0.198959	0.896904	-1.000000	-0.289644	-0.830434	-0.356376	-0.479785	-0.305079
12	-0.759798	-0.332863	-0.198959	0.896904	-1.000000	-0.289644	-0.830434	-0.356376	-0.479785	-0.305079	-0.761338
13	-0.332863	-0.198959	0.896904	-1.000000	-0.289644	-0.830434	-0.356376	-0.479785	-0.305079	-0.761338	-0.445912
14	-0.198959	0.896904	-1.000000	-0.289644	-0.830434	-0.356376	-0.479785	-0.305079	-0.761338	-0.445912	-0.677681

Tabla 7. Tabla que representa la capa de entrada de la red neuronal artificial creada.

En esta tabla se observa que las entradas son las columnas encabezadas como `var1(t-10)` a `(t-1)` y la variable de salida (lo que sería nuestra `y` de la función) será la última columna titulada como `var1(t)`.

Luego del procesamiento, y antes de crear la red neuronal artificial, se subdivide el conjunto de datos en `train` y en `test`. Algo importante a destacar en el procedimiento es que a diferencia de otros problemas en los que se puede hacer una mezcla de datos de entrada, es que en este caso importa mantener el orden en el que se alimentará la red neuronal artificial. Por tanto, se hace una subdivisión de las primeras 110 semanas consecutivas para el entrenamiento de la red y las siguientes 31 semanas para su validación. Esta es una proporción que se eligió de manera arbitraria pero que logro obtener buenos resultados y ser conveniente respecto a otras que se probaron, sin embargo no se puede precisar que esta sea la óptima.

Con lo anterior, se transformó la entrada en un arreglo con forma `(110, 1, 10)` esto en palabras simples significa algo del tipo “110 entradas con vectores de 1×10 ”. De esta manera la arquitectura de la red tiene las siguientes características:

- Entrada 10 inputs.
- 1 capa oculta con 10 neuronas.
- La salida será sólo 1.
- La función de activación utilizada es la tangente hiperbólica puesto que se utilizan valores entre -1 y 1.
- Se utilizará como optimizador “Adam” y métrica de pérdida (loss) Mean Absolute Error.
- Dada que la predicción es un valor continuo y no discreto, para calcular el Accuracy se utiliza Mean Squared Error y para saber si mejora con el entrenamiento este debería ir disminuyendo a medida que se realizan los EPOCHS, los cuales son la cantidad de veces que el modelo “aprende”, en este caso, la cantidad de EPOCHS utilizada es de 100 iteraciones.

Luego, se entrena y entregan los primeros resultados para ver cómo se comporta la red al cabo de las 100 épocas simuladas, se puede apreciar en la interfaz una reducción del valor de pérdida tanto

del set de entrenamiento como del de validación. Ahora se visualiza en el siguiente gráfico el conjunto de validación:

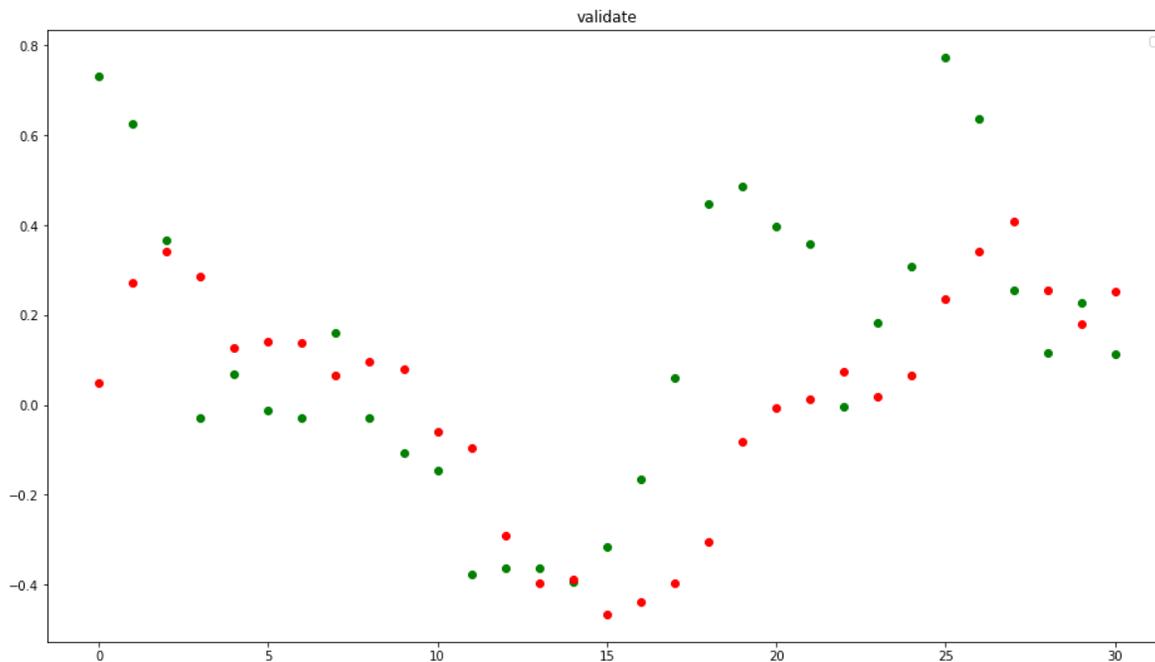


Gráfico 22. Valores predichos(rojo) vs valores del conjunto de datos de testeo(verde)

Se aprecia en el gráfico que los puntos verdes intentan aproximarse a los rojos, de manera que, a medida que están más cerca o superpuestos, es mejor. Se evidencia que al aumentar la cantidad de EPOCHS el resultado va mejorando.

Por otro lado, Se puede apreciar y comparar cómo disminuye el LOSS tanto en el conjunto de entrenamiento como en el de validación, esto tiene buenos indicios ya que se aprecia que el modelo está aprendiendo, y a su vez, pareciera no haber overfitting, pues las curvas de train y validate son distintas.

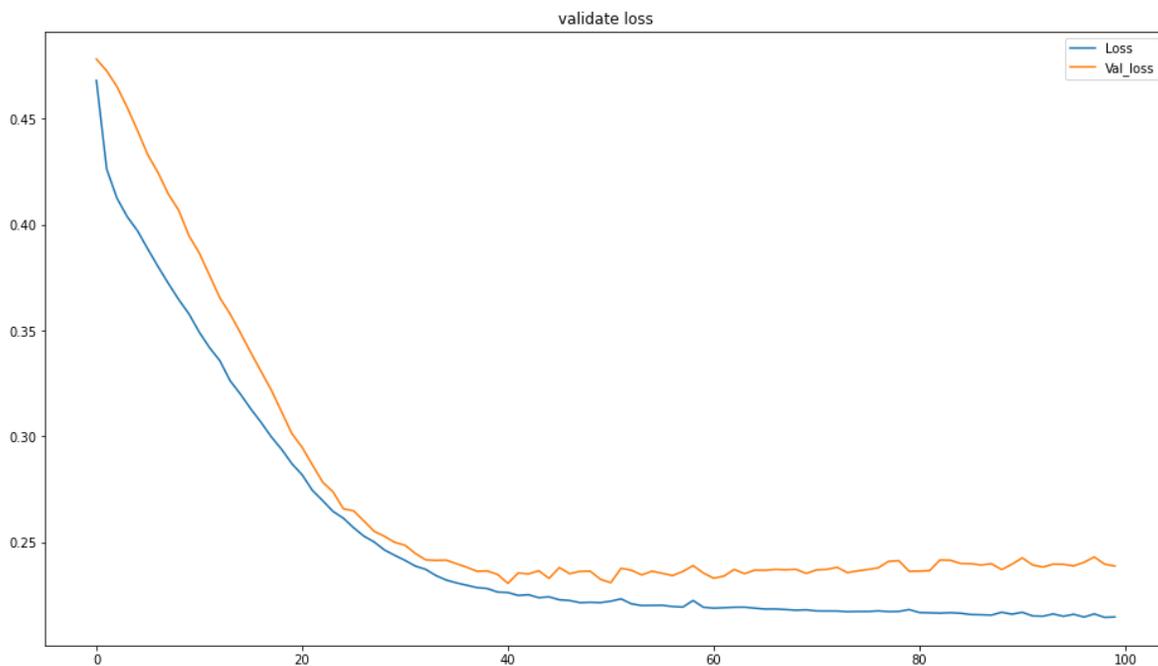


Gráfico 23. Función de pérdida vs función de pérdida del conjunto de testeo

Con esto, se pueden empezar a realizar predicciones para comparar con los valores reales, ya se tiene la red y se realizaron los 7 pasos del ML (colectar datos, preparar datos, elegir modelo, entrenar la máquina, evaluar, configurar parámetros y predicción o inferencia) por tanto se da el visto bueno y se prueba realizando predicciones, en este caso y para el resto de las categorías, se usarán datos de las últimas 102 semanas para predecir las semanas siguientes.

A continuación, se presenta una tabla que contienen los primeros 20 valores predichos vs los valores reales:

	Real	Predicción	Diferencia
0	90392	62425	27966
1	86070	71559	14511
2	75517	74395	1122
3	59255	72130	-12874
4	63231	65682	-2451
5	59994	66244	-6249
6	59311	66126	-6815
7	67054	63107	3947
8	59214	64370	-5156
9	56040	63761	-7721
10	54441	57975	-3533
11	45023	56549	-11525
12	45542	48541	-2998
13	45567	44251	1315
14	44346	44579	-232
15	47548	41399	6148

16	53677	42454	11222
17	62958	44179	18779
18	78771	48002	30769
19	80416	57078	23338

Tabla 8. Comparación entre los valores reales de la base de validación y las predicciones de la red neuronal artificial.

Visualizamos los resultados de la curva real versus la curva predicha por el modelo, en el siguiente gráfico:

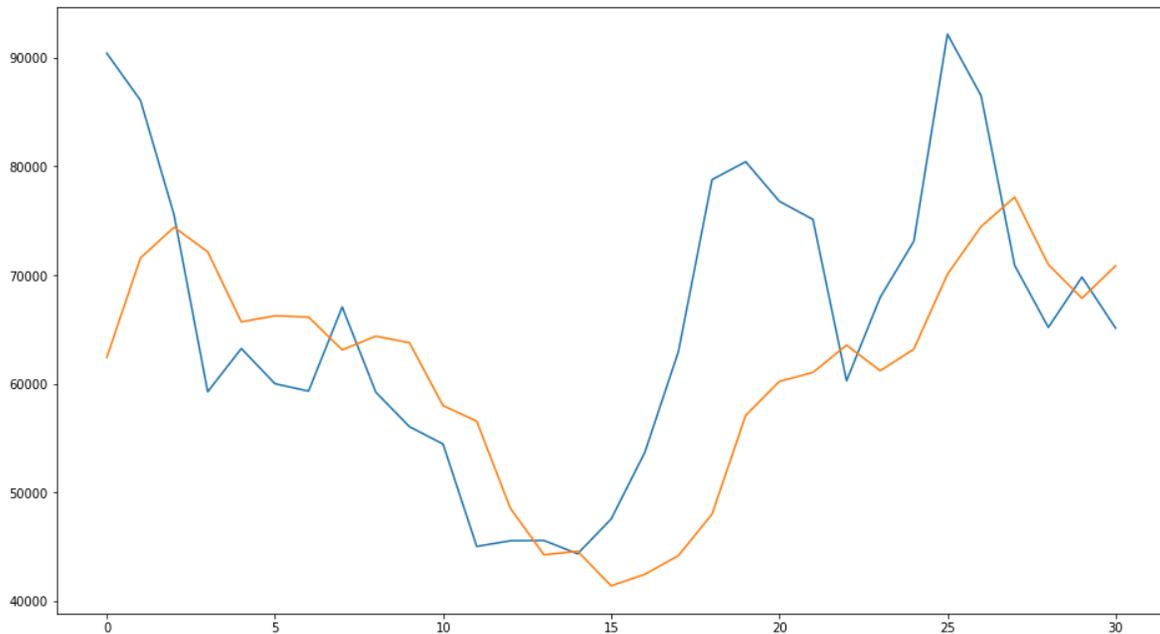


Gráfico 24. Predicción (naranja) vs valores reales (azul) del modelo de red neuronal artificial.

Ya con esto realizado, se procede a calcular el MAPE de esta predicción, obteniendo un valor de 14,04% para el caso de esta serie temporal, por lo que cumple con ser mejor que los métodos mostrados previo a este modelo y cumple con estar dentro del rango aceptable, es decir menor a un 18%. Cabe destacar que estos resultados son para esta categoría en particular, y sólo responden a presentar un ejemplo de la metodología que se siguió para realizar pronósticos y obtener los indicadores de MAPE por categoría de productos. El resto de los productos y categorías de los distintos sectores (cerdo, aves, procesados) se trató con la misma metodología cuyos resultados finales se encuentran en la sección de anexos de este informe.

8. Proyecto

8.1. Cadena de suministro de Agrosuper

El trabajo de título se desarrollará en la Gerencia de Cadena de Suministro (CDS) de Agrosuper compuesta por 124 colaboradores. En particular, en la Subgerencia de Planificación Comercial (PC) compuesta por 7 personas. En la siguiente figura se presenta el organigrama de la CDS y en mayor detalle la composición de profesionales que son parte de esta subgerencia.

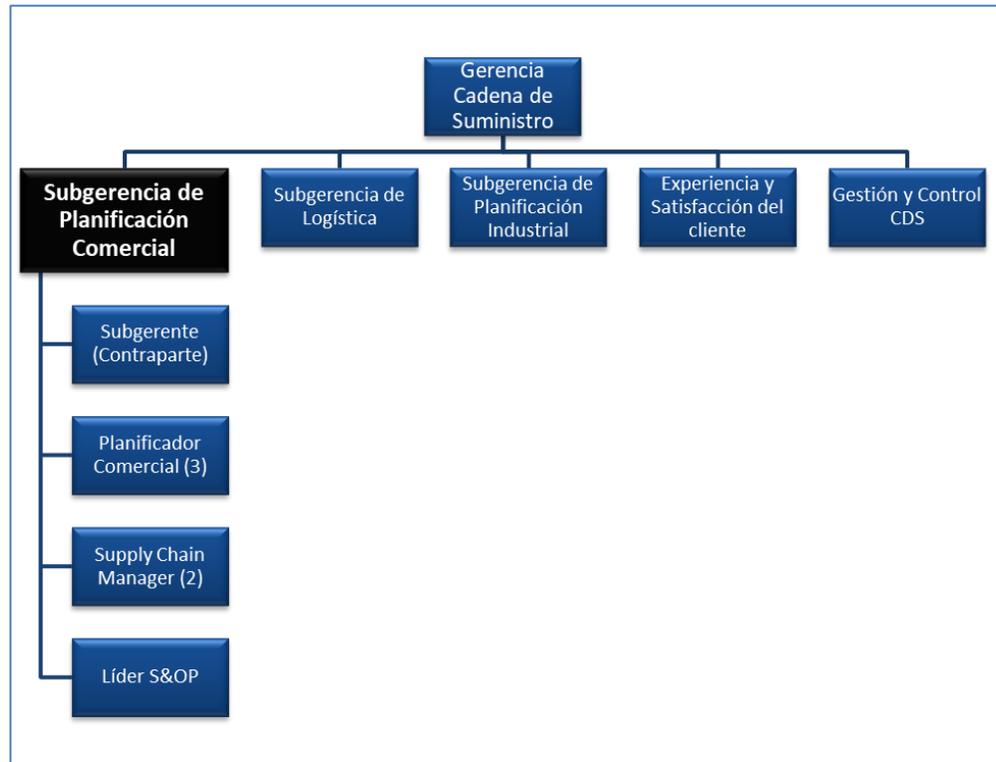


Figura 19. Organigrama de la CDS de Agrosuper y de la subgerencia de PC. Elaboración propia.

El propósito de esta Subgerencia es gobernar el proceso S&OP generando planes comerciales a mediano plazo, vale decir, en un horizonte de 1 a 3 meses, que conecten los objetivos del negocio con la satisfacción de clientes y consumidores. Además, deben liderar procesos en la Cadena de Suministro que movilicen la toma de decisiones ágiles para garantizar una correcta ejecución de estos planes comerciales tanto en el mercado nacional como en el internacional.

Los profesionales y las funciones que desempeñan en el área de trabajo son las siguientes:

- **S&OP Planner (3):** Liderar las instancias, roles y responsabilidades de los equipos. Generar un proceso de demanda táctico que permita detectar y capturar oportunidades rentables en ambos mercados para impactar positivamente el negocio.

- **Líder proceso S&OP (1):** Controlar la ejecución del proceso buscando la mejora continua, estandarización de buenas prácticas, y la adhesión/conexión del proceso con los clientes y consumidores.
- **Supply Chain Manager (2):** Asegurar la correcta ejecución del plan comercial. Coordinar y conectar a los equipos de la CDS con el proceso semanal del S&OP. Buscar a través del ciclo de demanda operativa mejoras al resultado del negocio en los diferentes mercados.

El solicitante del trabajo es el subgerente de PC, quien es el encargado de dirigir y liderar esta subgerencia y espera que este proyecto impacte directamente en mejoras de los planes comerciales que se generan mes a mes, además de garantizar un apoyo más robusto para la toma de decisiones relevantes del proceso de demanda.

8.2. Proceso de S&OP y su arquitectura

Se define el proceso de S&OP como aquel que sustenta la ejecución de la estrategia del negocio, constituye un método de planificación y comunicación, cuyo propósito es alinear y equilibrar los pronósticos comerciales con las restricciones industriales y logísticas, de manera tal que se cumplan los objetivos de rentabilidad esperados por la compañía.

Puede ser entendido como el proceso para desarrollar planes tácticos que provean a la Gerencia la habilidad para dirigir estratégicamente su negocio y así lograr ventajas competitivas sostenibles por medio de integrar planes de marketing y planes de producto y servicio nuevos y existentes, enfocados en el cliente con la gestión de la cadena de suministro. Lo importante de esta definición es que muestra que es un proceso y no un resultado y que en su ejecución alinea la estrategia con los planes tácticos, enfocándolo en la gestión de Supply Chain.

Al entender más profundamente esta definición y revisar todo lo que implica este proceso surge la inquietud si en la Agrosuper se usa adecuadamente y si finalmente logra el objetivo de ser una herramienta de alineación para lograr los resultados que busca la estrategia corporativa.

Para Agrosuper representa una herramienta de gestión que mejora la comunicación y planificación dentro de la empresa. Surge como solución a la vanguardia para problemas de falta de comunicación entre departamentos, lo cual termina afectando los procesos de alineamiento de objetivos, desfasando las estrategias e incrementando costos operacionales.

Frente a esta problemática, el S&OP funciona como un plan que involucra a todas las áreas pertinentes de una empresa e integra los planes de ventas dentro del proceso de administración de la cadena de suministros. Esto, en busca de la consecución de metas previamente establecidas, las cuales serán monitoreadas de manera periódica.

El método ha sido implementado en empresas grandes y medianas con el foco de conseguir beneficios tanto cualitativos como cuantitativos, dentro de los cuales se mencionan:

1. Mejora de la comunicación horizontal entre los departamentos de la empresa, logrando una sincronía especial entre las áreas de venta y logística.
2. Reducción de costos operativos al reducir el exceso de inventarios, los retrasos en la programación de producción y los gastos innecesarios de transporte.
3. Mayor control de inventario, reduciendo la suspensión de ventas a causa de insuficiencia de productos en puntos de venta o almacenes.
4. Procesos de gestión más eficientes al generar proyecciones y pronósticos que permiten detectar de manera anticipada posibles situaciones problemáticas u oportunidades positivas para el negocio.
5. Mejor control de desempeño y resultado de nuevos productos y promociones como parte del desarrollo de una estrategia global.
6. Si bien el S&OP está enfocado en una mejor gestión de las áreas de una empresa, para ser efectivo no debe dejar de enfocarse en el valor de los clientes y en la manera en que estos se conectan con los productos.

Cada ciclo se lleva a cabo en un período de dos meses, el foco del primer mes está en construir el plan comercial para el siguiente período a través de las 4 primeras etapas y el foco del segundo mes está en asegurar su ejecución, seguimiento y control de desviaciones a través de la última etapa. Las etapas del proceso son 5 como se muestra en la figura 20: generación de demanda; inputs productivos y stock; optimización y balance entre oferta y demanda; comunicación; y, por último, seguimiento y retroalimentación.

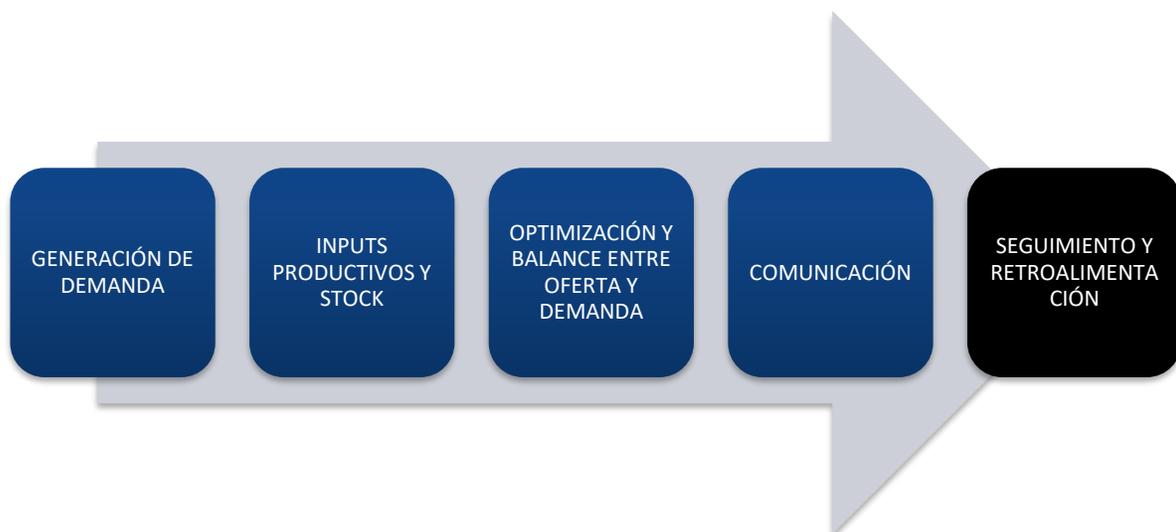


Figura 20. Ciclo de S&OP de la CDS de Agrosuper. Elaboración propia.

La planeación de ventas y operaciones (S&OP) compromete e involucra a varias áreas del negocio, con la finalidad de garantizar una mejor toma de decisiones y que estas trabajen de manera

colaborativa y permanentemente con la Subgerencia de PC a través de la repetición del ciclo, entre ellas se mencionan las siguientes:

- Planificación Industrial
- Ventas Nacionales
- Ventas Internacionales
- Logística
- Abastecimiento
- Negocios
- Revenue Management
- Producción Animal
- Plantas

8.3. Levantamiento de la situación actual

Para describir la situación actual del negocio se presenta la arquitectura de procesos definida en la metodología de la Ingeniería de Negocios. Se presenta la arquitectura del proceso de S&OP con las características del caso particular de la empresa modelando a través de los diagramas BPMN la situación actual, el funcionamiento, la dinámica de trabajo y las principales actividades que ejecutan cada una de las áreas involucradas y participantes del proceso presentadas previamente.

En la figura 21, se muestran las dos primeras etapas del proceso de S&OP (Modelamiento de Demanda y Oferta de Agrosuper) en dicha instancia se parte con la actualización del portafolio de productos de Agrosuper, para luego generar los pronósticos de demanda. Luego, la gerencia de ventas entrega su visión de la demanda futura y los inputs de promociones para finalmente confeccionar la Demanda única nacional conocida y validada por todas las áreas involucradas en este subproceso. Con la demanda confeccionada se actualiza la oferta de productos en donde se entrega principalmente por la subgerencia de planificación industrial la actualización de los inputs de producción animal, se ajustan y actualizan las restricciones productivas industriales por parte de las plantas de Agrosuper, para finalmente conocer las proyecciones de Stock disponibles para ser comercializadas. Es importante mencionar que este proyecto impactará a la etapa/subproceso de Generación de demanda.

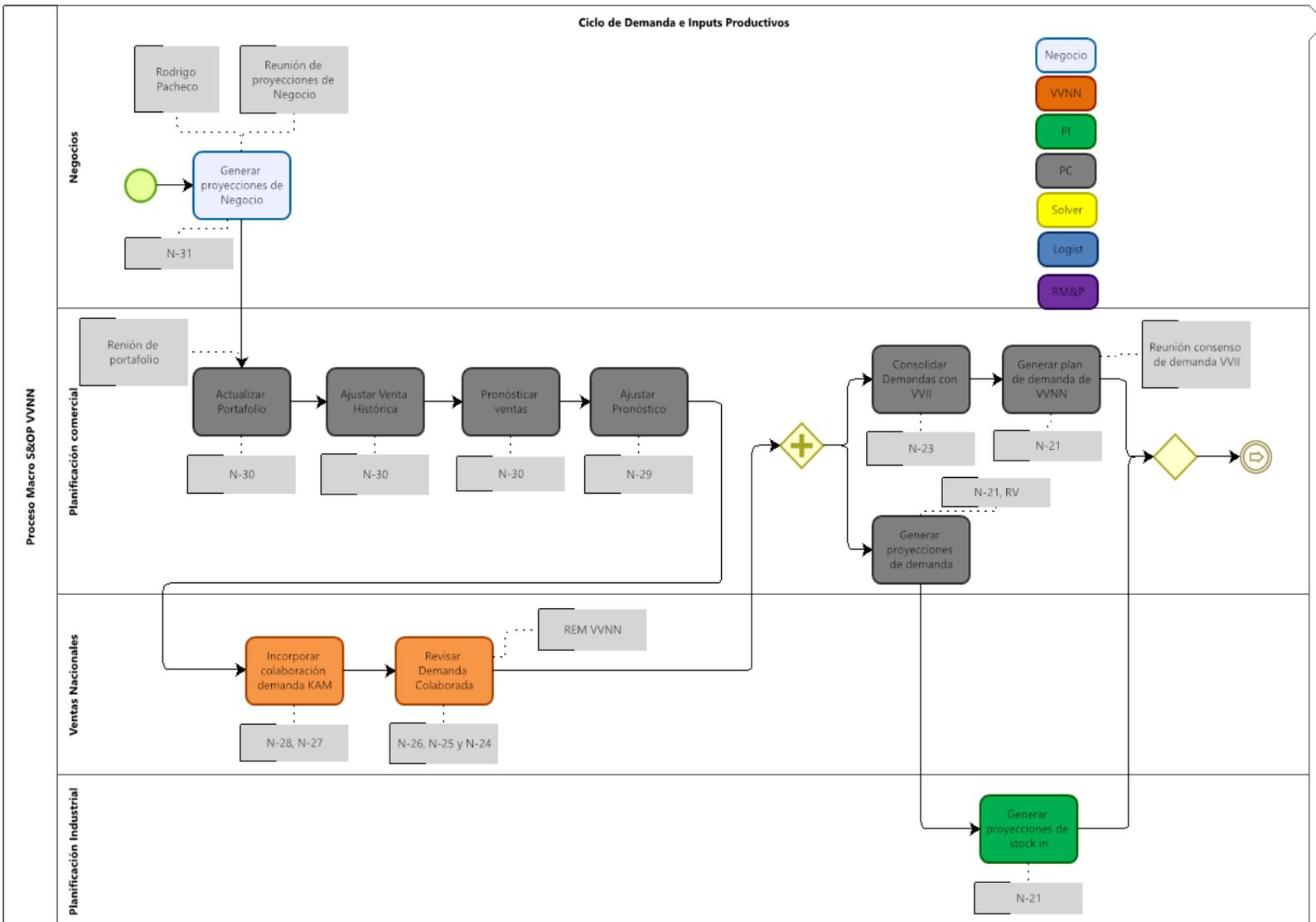


Figura 21. Etapas de Demanda e inputs productivos del proceso de S&OP. Elaboración propia.

Luego de haber modelado la demanda y oferta, se sigue la fase de optimización y equilibrio entre ambas variables, en esta etapa se revisan las estrategias, contratos y cotas comerciales. En primer lugar, se ejecuta un algoritmo que entrega la primera versión del plan comercial, con el resultado de este, se entrega a la subgerencia de Revenue y Management para que ellos calibren los precios de acuerdo con la cantidad de kilos disponibles, luego se actualizan las demandas y ofertas para nuevamente ejecutar el algoritmo de optimización (Solver). Con esto, se procede a modificar y actualizar los planes comerciales, en este periodo se hace una reunión denominada “Pre-volumen”

Y por último, se procede a comunicar y aprobar el plan por las comisiones del ejecutivo de la compañía, en donde participan distintos gerentes y subgerentes de Agrosuper. Con el plan S&OP ya aprobado se desagrega el plan por sucursal y canal de venta, para ser comunicado y subido a las plataformas de gestión de Agrosuper.

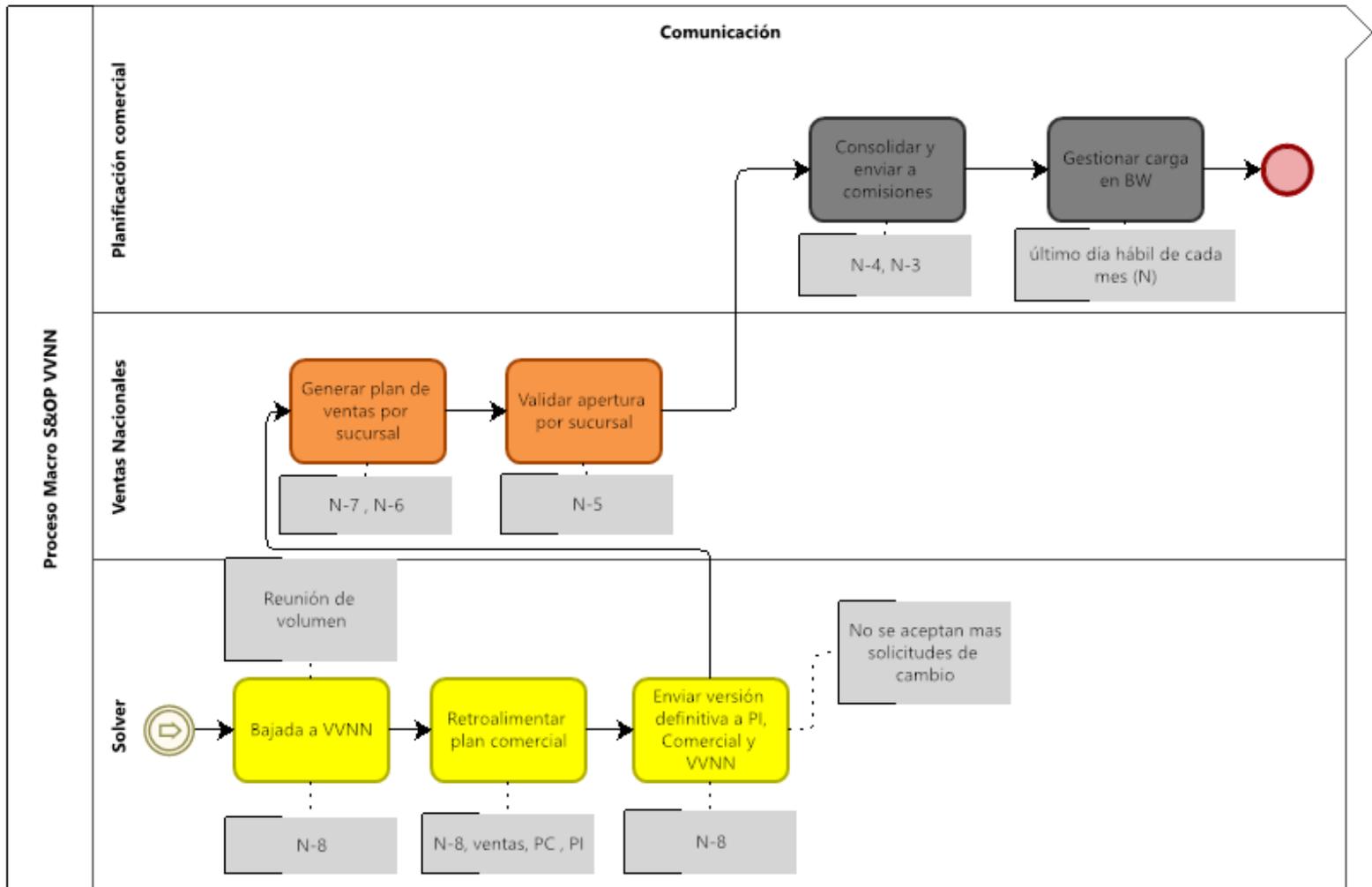


Figura 23. Etapa de Comunicación del proceso de S&OP. Elaboración propia.

Cabe destacar que la sigla N-X bajo cada tarea en los diagramas representa el día límite para cumplir con dicha tarea en el mes, es decir, cada actividad tiene un plazo establecido para una correcta ejecución del proceso y junto a esto también se tiene a los responsables de llevarlo a cabo.

8.4. Subproceso de generación de demanda

Dado que el proyecto de título se enmarca en este subproceso de generación de demanda del proceso general de S&OP, se hace necesario ahondar en mayor medida en la dinámica de este.

Proyecciones del negocio

Esta corresponde a la primera instancia del ciclo de demanda donde se comunican inputs cuantitativos y cualitativos de los mercados nacionales e internacionales que son relevantes para los negocios, por ejemplo, proyecciones de precios, variables y números macroeconómicos relevantes (tipo de cambio, inflación, entre otros). En dicha instancia se reúnen las áreas de venta, S&OP, Revenue y el área de Negocios como responsable, con una frecuencia mensual. El foco está en alinear todos los participantes del proceso. En la situación actual se realiza lo siguiente en dicha instancia:

1. Proyecciones del negocio
2. Tendencias Nacionales e internacionales
3. Volúmenes y precios mínimos - Nivel General

Actualización del portafolio

La segunda actividad de este subproceso es una instancia más técnica para preparar los sistemas y el portafolio de productos a planificar, en dicha instancia se analizan y diseñan estrategias/planes para productos en lanzamiento y para los productos a discontinuar, se analizan productos “espejo” para la demanda de productos similares/nuevos, se analizan formatos que no han tenido buenos resultados, se diseñan estrategias/planes de entrada para productos en lanzamiento y planes de salida para productos discontinuados, entre otras actividades técnicas. Se realiza una reunión con frecuencia mensual a cargo del equipo de S&OP cuyos asistentes son: Trade Marketing, Insumos, Datos Maestros, S&OP, Planificación Industrial, Gestión ventas, Desarrollo de productos y representantes de plantas (Abastecimiento). En la situación actual ocurre lo siguiente:

1. Incorporación de lanzamientos al sistema.
2. Creación de Productos nuevos en SAP IBP
3. Análisis de stock de productos ya discontinuados
4. Optimización de portafolio en SAP IBP
5. Confirmar Materias primas para lanzamientos

Baseline

Luego de tener el portafolio actualizado, se empieza a analizar la perspectiva cuantitativa de la demanda, tenemos la generación de predicciones de demanda a partir de modelos estadísticos puros tanto de análisis de series de tiempo, causalidad y simulación. Esta previsión estadística se ajusta en base a la segmentación de productos (ABC/XYZ) y tiene como input clave las demandas históricas. Representa una instancia de trabajo operativo de los planificadores comerciales

utilizando la herramienta SAP IBP y Excel, su frecuencia también es mensual. En la situación actual se realiza lo siguiente:

1. Revisar y validar el repositorio de datos históricos
2. Manejo de Outliers y BBDD
3. Segmentar el portafolio (ABC/XYZ; Volumen/variabilidad)
4. Asignar modelo estadístico según patrón de datos (Tendencia, estacionalidad, intermitencia)

El Baseline como su nombre lo indica, es la línea base de la historia de demandas, representa los pronósticos estadísticos puros o brutos, en esta instancia se corren distintos modelos estadísticos y mediante una funcionalidad de la herramienta de SAP IBP en donde el planificador comercial puede elegir el modelo que genera un menor MAPE y la herramienta lo hace de manera automática.

Segmentación de productos

Antes de continuar con este apartado, es necesario explicar la segmentación de productos que se hizo. Por regla de Pareto se hizo una segmentación de producto en base a la venta (volumen) histórico. Donde se define la clasificación ABC la cual se explica a continuación:

- **Materiales A:** Alto Volumen (Representa el 80% de la venta)
- **Materiales B:** Medio Volumen (Representa el 17% de la venta)
- **Materiales C:** Bajo Volumen (Representa el 3% de la venta)

Con lo anterior, se obtuvo como resultado la siguiente tabla:

Portafolio de productos										
Sector	AX	AY	AZ	BX	BY	BZ	CX	CY	CZ	Total
Cecina	16	10	10	6	16	10	3	8	21	100
Cerdo	16	39	80	13	25	82	8	13	93	369
Elaborado	10	16	18	2	10	19	4	2	23	104
Pavo	10	4	20	3	11	18	1	4	18	89
Pollo	10	13	41	4	14	32	4	13	43	174
Total	62	82	169	28	76	161	20	40	198	836

Tabla 9. Segmentación del portafolio de productos según volumen. Realizada por el equipo de S&OP.

Para enfocar esfuerzos se clasifico la variabilidad de los materiales en base a su comportamiento histórico y su varianza muestral. La clasificación XYZ se define con los siguientes parámetros:

- **Materiales X:** Baja Variabilidad. Desviación entre 0% y 30% (**X**)
- **Materiales Y:** Media Variabilidad. Desviación entre 31% y 50% (**Y**)
- **Materiales Z:** Alta Variabilidad. Desviación entre 51% y 100% (**Z**)

De esta manera, se obtuvo la siguiente tabla resultante:

Sector	AX	AY	AZ	BX	BY	BZ	CX	CY	CZ
Cecina	16%	10%	10%	6%	16%	10%	3%	8%	21%
Cerdo	4%	11%	22%	4%	7%	22%	2%	4%	25%
Elaborado	10%	15%	17%	2%	10%	18%	4%	2%	22%
Pavo	11%	4%	22%	3%	12%	20%	1%	4%	20%
Pollo	6%	7%	24%	2%	8%	18%	2%	7%	25%
Sugerencia	Forecast		Col	Forecast		Col	Forecast		Col

Tabla 10. Segmentación del portafolio de productos según variabilidad de ventas. Realizada por el equipo de S&OP.

En las combinaciones AZ;BZ;CZ es importante un análisis cualitativo, dada la alta variabilidad de los productos, ahí se debe hacer un análisis complejo de colaboración. Mientras que para la planificación de las demás categorías se definió que estas serán planificadas por medio de un pronóstico de los modelos estadísticos.

Forecast Inicial

Ajuste o manipulación del pronóstico estadístico inicial (baseline) en base a un análisis cualitativo el cual incluye lineamientos de proyecciones, conocimiento del mercado, estrategias y su experiencia, y en base a un análisis cuantitativo que se enfoca en observar las demandas más recientes del negocio. Los responsables de esta instancia es el equipo de S&OP y se hace de manera mensual, en estricto rigor esta instancia representa horas de trabajo operativo. En relación con la dinámica de la situación actual se comenta lo siguiente:

1. Análisis del baseline (volumen y precio) grandes números (A,B / X)
2. Corrección de volúmenes y precios. (Análisis situacional)
3. Análisis de productos nuevos (mínimos buscados)

Inputs de venta

Luego sigue la colaboración del input táctico que entrega la subgerencia de ventas, esta incorpora el potencial de los canales a través de contribuciones por gestión, promociones y lanzamientos. Esta etapa se compone de dos instancias (Pre-REM y REM) los responsables son el área de ventas y su frecuencia también es mensual. En la situación actual se opera de la siguiente manera:

1. Análisis de potencial de ventas
2. Análisis de demanda y kpis claves de ejecución
3. Inclusión de productos nuevos
4. Análisis promocional y de estacionalidades

Plan Demanda Consensuada

Por último, se tiene una instancia de definición de una demanda única, conocida y validada por todas las áreas involucradas en modelar la demanda, con gerentes de canales de venta que entregan el input de demanda final (kilos y precios) que será enviado al proceso siguiente de optimización. Cabe destacar que este valor da origen a la “demanda irrestricta”, la cual representa la proyección de demanda sin restricciones logísticas o productivas de lo que se espera que clientes y consumidores demanden. Dicha instancia, consta de una reunión cuya responsabilidad recae en el equipo de planificación comercial. Los asistentes son los gerentes y subgerentes de los distintos canales de Agrosuper y su frecuencia es mensual. En la situación actual se realiza lo siguiente:

1. Resumen del Plan a nivel de canal/segmentación principales Niveles 2
2. Revisión de las colaboraciones de demandas irrestrictas
3. Validación de volúmenes y precios

Finalmente, se llega al **plan consensuado de demanda**, y este plan representa el término del subproceso de generación de demanda. En los siguientes gráficos se muestran los indicadores de MAPE de los distintos negocios, es decir, la diferencia porcentual entre las proyecciones de cada una de estas variables respecto a la demanda real que se obtiene posterior a la confección de estos planes.

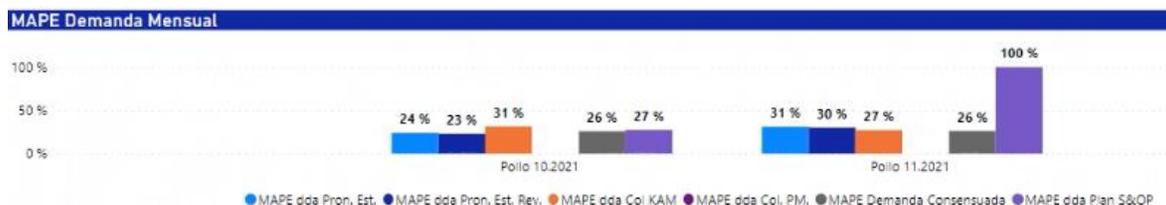


Gráfico 25. MAPE de demanda mensual del sector pollo en los meses de Octubre y Noviembre del año 2021 en el mercado nacional. Fuente: Agrosuper.

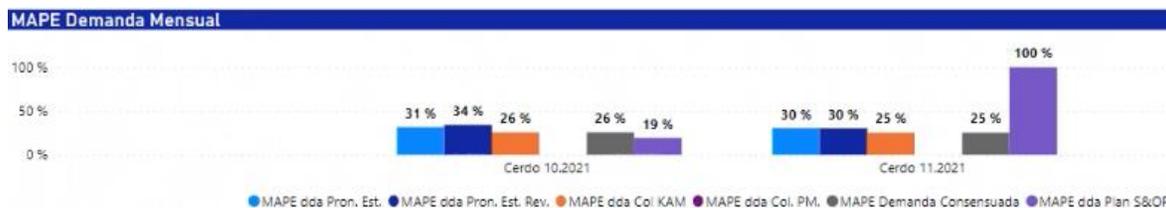


Gráfico 26. MAPE de demanda mensual del sector cerdo en los meses de Octubre y Noviembre del año 2021 en el mercado nacional. Fuente: Agrosuper.

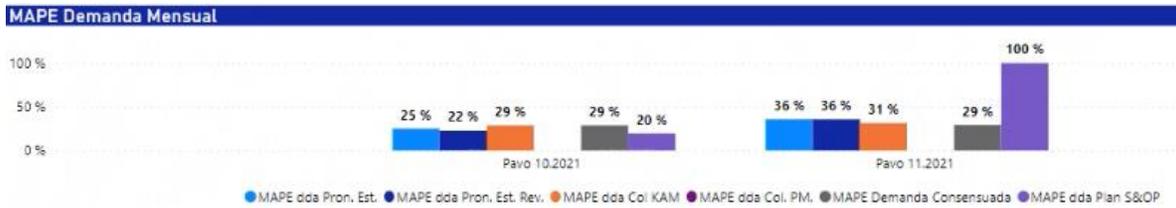


Gráfico 27. MAPE de demanda mensual del sector pavo en los meses de Octubre y Noviembre del año 2021 en el mercado nacional. Fuente: Agrosuper.

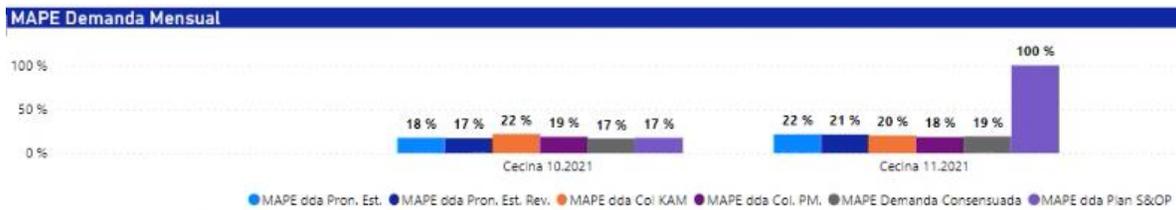


Gráfico 28. MAPE de demanda mensual del sector cecina en los meses de Octubre y Noviembre del año 2021 en el mercado nacional. Fuente: Agrosuper.

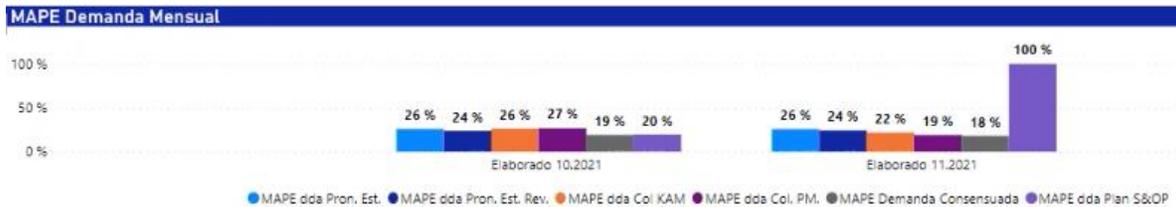


Gráfico 29. MAPE de demanda mensual del sector pollo en los meses de Octubre y Noviembre del año 2021 en el mercado nacional. Fuente: Agrosuper.

A continuación, se muestra en la figura 24 el avance de este plan de demanda y en la tabla 11 el total de kilogramos que se pronostican en un mes normal por sector en cada una de las instancias del proceso de estimación de demanda, con el fin de dimensionar el nivel de operación.

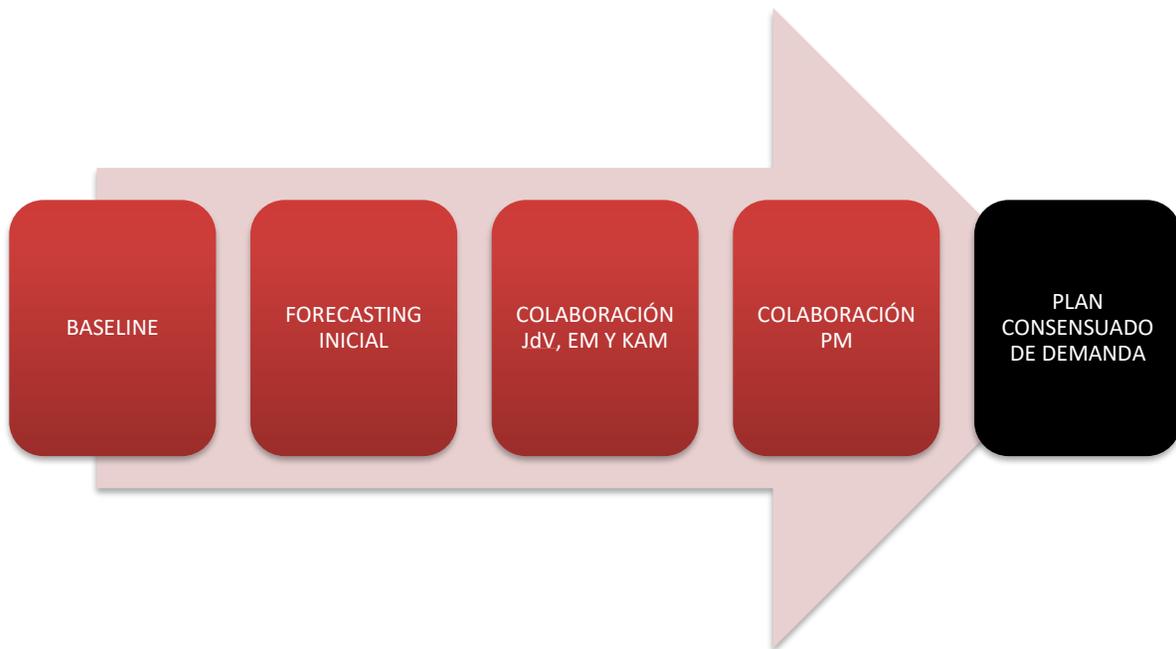


Figura 24. Evolución del plan de demanda. Elaboración propia.

	Cecina	Cerdo	Elaborado	Hortalizas y Frutas	Pavo	Pollo	Salmon	Vacuno
Base Line Kg	2.514.678	17.730.303	2.536.009	38.622.766	18.472.050	291.326.672	170.070	627
Forecast Inicial Kg	2.514.678	19.082.553	2.536.009	38.622.766	3.255.108	31.389.079	170.070	627
Colaboración JdV, KAM y EM Kg	2.619.715	18.821.191	2.213.276	38.528.857	3.334.966	28.872.328	152.234	627
Colaboración PM Kg	2.701.527	18.821.191	2.213.276	38.528.857	3.334.966	28.872.328	152.234	627
Plan Consensuado de Demanda Kg	2.770.869	19.135.053	2.011.326	38.528.857	3.358.416	29.071.202	152.234	627

Tabla 11. Valores de la evolución del plan de demanda para todos los sectores de Agrosuper. Fuente: Agrosuper.

8.5. SAP IBP

La tecnología utilizada para este proceso es SAP IBP, esta herramienta permite dar una completa trazabilidad al proceso de S&OP y combina la supervisión de la cadena de aprovisionamiento, la planificación de ventas y operaciones, la gestión de demanda, la planificación de inventario y la planificación de aprovisionamiento en una solución integral que, además, admite la integración de datos de sistemas externos. SAP IBP proporciona aplicaciones que facilitan: [10]

1. Compensación de la demanda y el aprovisionamiento.
2. Armonización de las planificaciones en todas las funciones empresariales a través de la visibilidad organizacional y de la alineación.
3. Planificación responsable con algoritmos avanzados, optimización y planificación de escenarios de simulación.

4. Planificación de cadena de aprovisionamiento en función de la demanda basada en el Demand Sensing, el análisis y la previsión predictiva.
5. Optimización del inventario de varias fases y niveles de servicio al cliente mejorados.
6. Visibilidad de la cadena logística integral con supervisión en tiempo real, analíticas, alertas y gestión de excepción.
7. Creación de contingentes y planes de aprovisionamiento, y habilidad para responder a las modificaciones con simulaciones y análisis cuello de botella.

La siguiente figura ilustra el proceso de alto nivel y, a continuación, se explican las distintas fases del proceso asociadas a la demanda que esta herramienta ofrece:

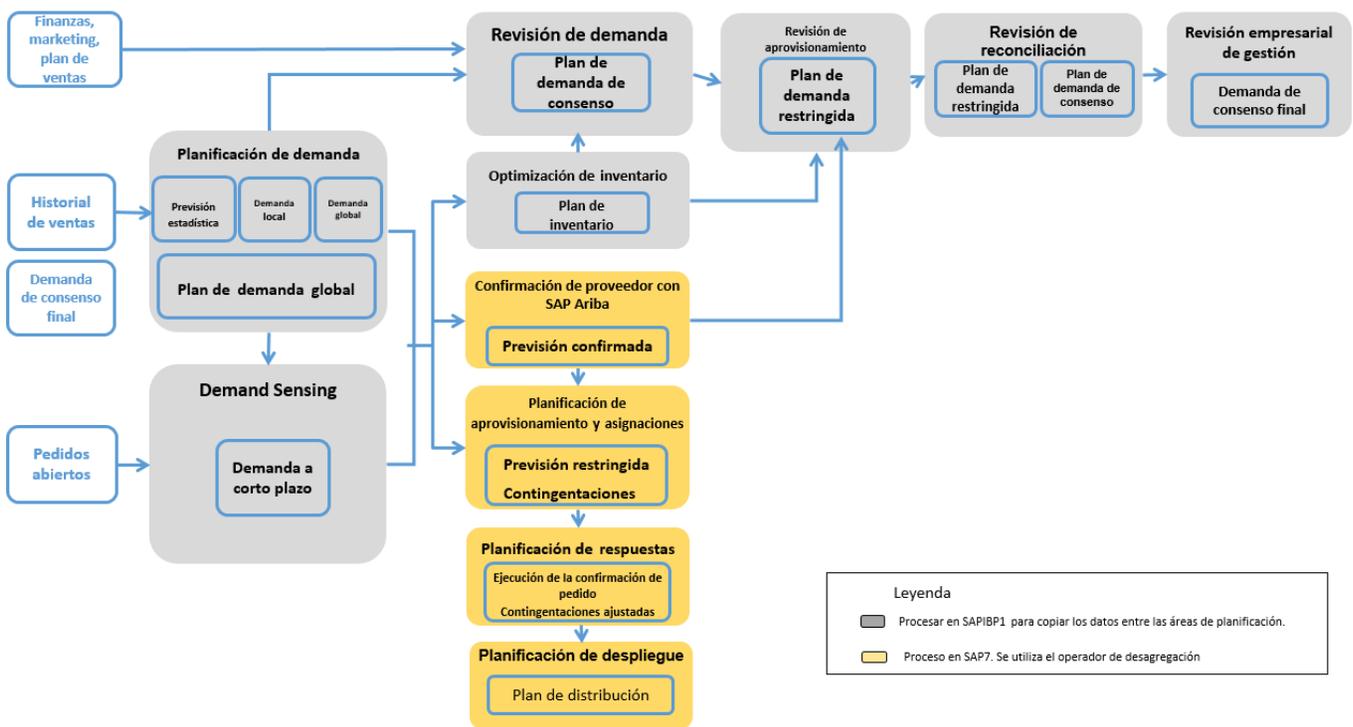


Figura 25. Proceso de planificación integrado con el área de planificación unificada usando SAP IBP.
Fuente: Agrosuper y SAP IBP.

Planificación de demanda

En la fase de planificación de la demanda, el planificador de demanda planifica los próximos meses. Esto se realiza en ciclos semanales e incluye las siguientes tareas:

1. Crear previsión estadística

El experto en el proceso de planificación de la demanda crea la previsión estadística basada en los datos históricos de ventas.

2. Crear un plan de demanda local

Basado en la previsión estadística, el planificador de la demanda local crea el plan de demanda local, normalmente para un producto de ubicación específica o un grupo de productos.

3. Crear un plan de demanda global

Basado en el plan de demanda local y el consenso final de demanda definido en las ventas previas y el ciclo de planificación de operaciones, el planificador de demanda global crea el plan de demanda global.

4. Demand Sensing

En la fase de Demand Sensing, el sistema crea la demanda a corto plazo de la siguiente manera: Basado en la cantidad de pedido futura, la cantidad confirmada, los datos de cantidad entregada de SAP ERP, y el plan de demanda global definido en la planificación de la demanda, la demanda a corto plazo se calcula diariamente. La demanda a corto plazo es una entrada para crear el plan de demanda combinado.

5. Revisión de demanda

En la fase de revisión de demanda, el planificador de demanda crea un plan de demanda de consenso completo en el horizonte de planificación de medio a largo plazo basado en el plan de demanda global mensual. El plan también tiene en cuenta las ventas y las entradas de marketing y asegura que se cumplan los objetivos financieros.

6. Integrar los datos de entrada de fuentes externas (como SAP ERP)

Algunos de los datos utilizados en los procesos de planificación de ventas y operaciones se tienen que recuperar de fuentes externas. Ejemplos de estos datos son el plan de marketing, el plan financiero y el plan de ventas.

7. Crear un plan de demanda de consenso

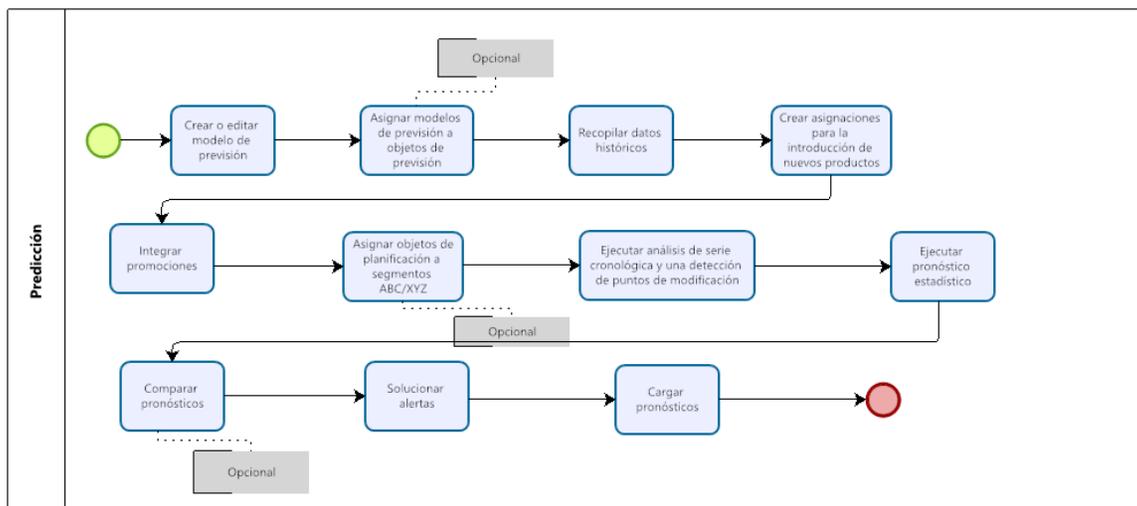
Durante la revisión de demanda, los responsables de ventas, finanzas, marketing y planificación de demanda definen un plan de demanda de consenso, mediante ratios de entrada derivados de la planificación de esta, de la optimización de inventario y de las fuentes externas.

8. Optimización de inventario

La optimización de stock se realiza semanalmente y utiliza la salida combinada de los procesos de planificación de demanda y de Demand Sensing para crear un plan de stock. La optimización de stock calcula los objetivos de stock para cada material en cada ubicación en una cadena logística, considera y compensa los imprevistos en las previsiones de demanda, el aprovisionamiento temporal y la cantidad de aprovisionamiento. El plan de stock sirve como entrada a los procesos de planificación de ventas y operaciones.

La planificación de la demanda ofrece varias herramientas para generar previsiones para diferentes escenarios o para determinadas partes de la empresa. La utilización de datos históricos y de algoritmos estadísticos con base científica permite mejorar la precisión de las previsiones de ingresos, alinear los niveles de inventario con los cambios de la demanda predecible y aumentar la rentabilidad de un canal o producto determinado.

El siguiente diagrama muestra los pasos obligatorios y opcionales que juntos comprenden el proceso de realizar una predicción de demanda.



Powered by
bizagi
Modeler

Figura 26. Proceso para realizar una predicción estadística con la herramienta de SAP IBP. Fuente: Agrosuper.

La previsión estadística es realizada por los planificadores comerciales, quienes pueden crear y editar modelos de previsión en la aplicación, además de gestionar modelos de previsión para lograr una variedad de resultados de previsión. Luego, pueden asignar sus modelos de previsión a varios objetos de planificación. Entre estos modelos se mencionan los siguientes:

- Alisamiento exponencial sencillo del porcentaje de respuesta dinámica
- Auto-ARIMA/SARIMA
- Auto-ARIMAX/SARIMAX

- Alisamiento exponencial automático
- Alisamiento exponencial lineal de Brown
- Copiar períodos pasados
- Método Croston
- Demand Sensing (completo)
- Demand Sensing (actualización)
- Alisamiento exponencial doble
- Potenciación del gradiente de árboles de decisiones
- Regresión lineal múltiple
- Regresión lineal temporal
- Promedio simple
- Valor medio variable simple
- Alisamiento exponencial sencillo
- Alisamiento exponencial triple
- Media ponderada
- Valor medio variable ponderado

Se hizo una comparación entre estos métodos, por lo que a continuación, en la siguiente tabla se presentan las principales funcionalidades que tienen estos modelos provistos por SAP IBP, vale decir, que la X representa que el modelo si puede gestionar dicha componente de la serie temporal en los pronósticos.

Es importante aclarar que se define la tendencia como el comportamiento o movimiento a largo plazo; la estacionalidad es la variación periódica y predecible de la misma serie temporal con un periodo inferior o igual a un año, y las series cronológicas intermitentes representan series de tiempo que entre sus valores existen términos positivos y valores que son nulos por falta de ventas u otros factores.

Algoritmo	Puede gestionar la tendencia en la serie cronológica	Puede gestionar la estacionalidad en la serie cronológica	Puede gestionar series cronológicas intermitentes
Alisamiento exponencial sencillo del porcentaje de respuesta dinámica	X		
ARIMA automático	X		X
SARIMA Automático	X	X	X
ARIMAX automático	X	X	X
SARIMAX automático	X	X	X
Alisamiento exponencial automático	X	X	X
Alisamiento exponencial lineal de Brown	X		
Copiar períodos pasados	X	X	X
Método Croston			X
Alisamiento exponencial doble	X		
Gradient boosting de árboles de decisiones		X	X
Regresión lineal múltiple	X	X	X
Regresión lineal estacional sin la opción Considerar estacionalidad seleccionada	X		X
Regresión lineal temporal	X	X	X
Promedio simple			X
Promedio variable simple			X
Alisamiento exponencial sencillo			X
Alisamiento exponencial triple	X	X	
Promedio ponderado	X	X	X
Promedio variable ponderada	X	X	X

Tabla 12. Comparación entre las funcionalidades de los métodos predictivos provistos por SAP IBP.
Elaboración propia.

Cabe destacar que algunos de estos modelos son extensiones funcionales de otros, en este trabajo de título se entregó en los apartados anteriores perspectiva teórica y práctica de lo que hacen la mayoría de estos, en particular el modelo de Auto ARIMA/SARIMA, regresiones lineales, suavizaciones exponenciales, entre otros. La mayoría representan métodos estadísticos tradicionales.

9. Rediseño o mejora del subproceso

El marco de referencia será el rediseño de procesos. Para ello, se plantean las variables para la dirección del cambio, el detalle de la mejora de procesos, las lógicas de negocio y el alineamiento del proyecto con la estrategia de la empresa.

A modo de contextualización se hizo una revisión de la etapa inicial del proceso S&OP “Generación de demanda”, con la colaboración y participación de distintos agentes del equipo de planificación comercial y del área de ventas de Agrosuper, en donde se identificaron oportunidades de mejora dentro de las instancias y actividades que la componen.

Para clasificar el nivel de impacto se consideró una matriz que mide el esfuerzo en tiempo y una valorización económica preliminar. Con lo anterior, quedan los siguientes criterios que pueden observarse en la siguiente figura:

	+ 4 ciclos	3 ciclos	2 ciclos
Tiempo	ALTA	MEDIA	BAJA
Inversión	ALTA	MEDIA	BAJA
	\$ +1 Millón	\$ -1 Millón	\$ 0

Figura 27. Matriz que mide el nivel de impacto en la mejora del proceso de generación de demanda, medido en tiempo e inversión económica. Elaboración propia.

Donde se explicitan los niveles de impacto a continuación:

- **Bajo (B):** Significa que no requiere inversión (0 pesos chilenos) ni consultoría, solo horas internas de trabajo operativo y tiempo en implementación será de 2 ciclos, o bien dos meses.
- **Medio (M):** Se requiere la colaboración de más de un área para solucionar, hay inversión económica moderada (máximo 1 millón de pesos chilenos) y el tiempo de implementación será menor a 4 ciclos.
- **Alto (A):** Se requiere inversión (sobre 1 millón de pesos chilenos), capacitaciones internas y externas, sin plazo máximo o mínimo de implementación.

9.1. Direcciones de cambio

Para dar un marco de referencia que permita un rediseño del proceso a partir de la estrategia, modelo de negocio, arquitectura y situación actual, se plantean las variables asociadas a las innovaciones que se pueden efectuar en el proceso.

Se resume la estrategia, modelo de negocio y arquitectura del caso a continuación:

- El posicionamiento estratégico está dado por la diferenciación. Esto es, la participación con bajos costos, pero con características que hacen diferenciar a la empresa de sus competidores.
- El modelo de negocio consiste en crear valor a los accionistas a través de servicios de excelencia en el ámbito de la ingeniería, entregando soluciones integrales e innovadoras que agreguen valor a los proyectos de la cadena de suministro y Agrosuper en general.
- La arquitectura corresponde a la cadena de valor, presentada en los diagramas de la situación actual.

Dicho lo anterior, en este apartado se explican los principales cambios que se pretende realizar para las variables de Mantenimiento Consolidado de Estados, Anticipación, Prácticas de Trabajo, Coordinación y Apoyo Computacional.

Sobre la variable de Asignación de responsabilidades, si bien no es el foco principal del proyecto, en todos los procesos que se rediseñan se contempla la definición de responsables claros que, en general, son los mismos trabajadores actuales de la empresa.

Sobre la variable de Coordinación, se incluyen ciertas prácticas específicas para coordinar el área de ventas con la de planificación y también las diferentes áreas productivas.

Sobre la variable de mantenimiento consolidado de estados, debido a los problemas que se detectaron, se puede notar que la base para poder mejorar el proceso de demanda de la cadena de suministro tiene relación con la disponibilidad de información sobre los procesos y su estado, y según los diagnósticos la empresa tiene un buen resultado. Además, la información que se genere de los procesos debe ser confiable, es decir, no debe contener errores y debe ser de fácil uso, es decir, que el procesamiento para darle utilidad a la información no requiera demasiado tiempo o complejidad. Se debe considerar que la variable de mantenimiento de estados depende en gran medida de la variable apoyo computacional.

La Anticipación es otra variable de cambio relevante en el rediseño, sobre todo para la planificación de ventas, de la producción, abastecimiento, entre otros procesos involucrados en el modelo de S&OP. Uno de los cambios principales sobre esta variable consiste en reforzar la elaboración de los pronósticos de demanda mediante la incorporación de este nuevo método (RNA), la que permita anticiparse de mejor forma a los requerimientos que tendrán los clientes y programar la producción acorde a dicha estimación con el propósito principal de disminuir las ventas perdidas. La estimación de demanda también es clave para definir los pedidos y el abastecimiento de materias primas.

La variable de cambio de prácticas de trabajo tiene dos modificaciones fundamentales, que se refieren a las prácticas necesarias para realizar la mantención de estados y por otra parte la formalización de las prácticas que se realizan para llevar a cabo los procesos y actividades específicas que lo componen.

Para los procesos que cuentan con un mayor grado de formalización, y en los cuales se realizan registros en las plataformas, se definen las prácticas de trabajo específicas para pasar a realizarlos y generar más valor para las distintas áreas involucradas, lo que incluye el uso específico que se le da al software que sea seleccionado para definir el rediseño.

Por otra parte, hay otras prácticas de trabajo cuyo principal cambio tiene que ver con su formalización, es decir, plantear de forma precisa de qué manera se realizará y en caso de que lo incluya, bajo qué reglas se tomarán las decisiones.

Muchos de los cambios que se proponen para las variables recién nombradas, dependen en gran medida de tener un apoyo computacional que los sustente, ya que de lo contrario costaría mucho trabajo obtener y usar la información que se genera durante los procesos para alcanzar los objetivos planteados, por lo que esta variable es una de las más importantes del rediseño.

El software que se utilizará en el rediseño debería poder hacer las modificaciones necesarias sobre éste, para que se adapte a la estructura y procesos de la empresa y que el rediseño no se vea limitado por éste. Además, el software elegido para apoyar el rediseño debe permitir el procesamiento de datos e incluso la automatización de determinadas tareas. El apoyo computacional también debe ayudar en la comunicación entre las distintas áreas de la empresa, lo cual se debe realizar mediante disponibilidad en línea de la información. Esto permitirá, entre otras cosas, comunicar a miembros que se encuentren en distintas áreas de la empresa.

Principales cambios en los procesos

Se muestran a continuación los cambios principales que propone el rediseño, respecto a los procesos que existen actualmente en la empresa, los cuales se pueden ver en la siguiente tabla:

	Variable de cambio	Nivel de impacto	Principales cambios
Proyecciones del negocio	Coordinación	Bajo	Área de negocios sugiere precios mínimos y en los casos necesarios algún tipo de volumen que permita alinear el proceso con la estrategia de negocio para construir planes comerciales trimestrales
	Prácticas de trabajo	Bajo	Negocios comunica situación futura de materiales foco (tendencias) y dar énfasis en los nuevos lanzamientos de productos
	Coordinación	Bajo	Negocios indica holguras productivas, tendencias de precios y MAPE/Fillrate de productos críticos para focalizar el proceso de confección de demanda
	Prácticas de trabajo	Bajo	Negocios da apertura a lineamientos/estrategia para los diferentes canales de venta y mercados relevantes
Actualización del portafolio	Prácticas de trabajo	Medio	Confirmar los lanzamientos con casos de negocios adjuntos y validar insumos productivos
	Coordinación	Medio	Comunicar a los equipos de logística, abastecimiento, transporte y representantes de plantas productivas para incorporarlos al proceso
	Prácticas de trabajo	Medio	Optimizar portafolio incluyendo los planes de salida para productos a discontinuar y situación de insumos productivos
	Prácticas de trabajo	Medio	Dar seguimiento a los productos en lanzamiento, en base a criterios de éxito/fracaso
	Apoyo Computacional	Bajo	Actualizar y habilitar segmentación de productos en SAP IBP de acuerdo a volumen y precio
Baseline	Apoyo Computacional	Alto	Diseñar y ajustar modelos estadísticos predictivos, encontrar el mejor modelo para cada tipo de segmentación ABC/XYZ
	Anticipación	Alto	Entender teóricamente los modelos predictivos mediante capacitaciones y su aplicación según segmentación, se espera una curva de aprendizaje por parte de los planificadores comerciales
	Prácticas de trabajo	Medio	Revisar detalladamente resultados estadísticos generados por los modelos y potenciar el entregable a las demás áreas involucradas
	Anticipación	Alto	Generar modelos de pronóstico para precios
Forecast Inicial	Prácticas de trabajo	Medio	Revisar exhaustivamente los productos en precio y volumen en todos los segmentos
	Prácticas de trabajo	Alto	Realizar ajustes de precios en SAP IBP
	Mantención consolidada	Alto	Seguimiento y control del MAPE como base para mejorar forecast de demanda
	Anticipación	Alto	Integrar la perspectiva de los modelos de aprendizaje automático (RNA) a través de una colaboración que se hará a nivel 2 de los distintos materiales y sectores comercializados
	Coordinación	Alto	Separar focos de responsabilidad en base a segmentación ABC/XY
Inputs de venta	Prácticas de trabajo	Alto	Realizar colaboración de productos importados
	Coordinación	Medio	Colaborar a nivel segmentación y medir volúmenes colaborados según parámetros sugeridos en kickoff
	Mantención consolidada	Bajo	Estandarizar variables base para colaborar, en relación con la cantidad de meses, venta o demanda, fillrate o MAPE, IDA.
	Anticipación	Medio	Optimizar esfuerzos destinados a colaborar, utilizando forecast sugerido según segmentación con foco en productos ABC/Z
	Coordinación	Medio	Incluir Revenue como ente validador de precios colaborados por ventas
	Apoyo Computacional	Medio	Mantener SAP IBP operativo al 100% sin detalles técnicos
	Coordinación	Alto	Fundamentar espíritu de colaboración y realizar retroalimentación de la calidad de las colaboraciones, explicar grandes cambios según forecast
Plan demanda Com:	Prácticas de trabajo	Bajo	Analizar Baseline, Forecast vs colaboración de ventas, levantando oportunidades de mejora
	Prácticas de trabajo	Medio	Levantar holguras productivas para intensionar la demanda, resaltando el espíritu irrestricto de la demanda
	Coordinación	Medio	Focalizar productos de mayor volumen de venta, entre ellos envasados, lanzamientos, entre otros.

Tabla 13. Direcciones de cambio para el rediseño propuesto. Elaboración propia.

Con estos cambios y propuesta de mejora se espera:

- Reforzar los pilares del S&OP (Proceso, Herramientas, Gobierno, Roles) dentro de la etapa de **construcción de demanda**.
- Maximizar la calidad de los inputs mediante el **uso de herramientas y el trabajo colaborativo** de todos los equipos.
- Construir una demanda irrestricta de clase mundial.
- Mejorar las estimaciones de demanda medidas por el indicador de MAPE especialmente en.

Aún existen mejoras por implementar como lo son:

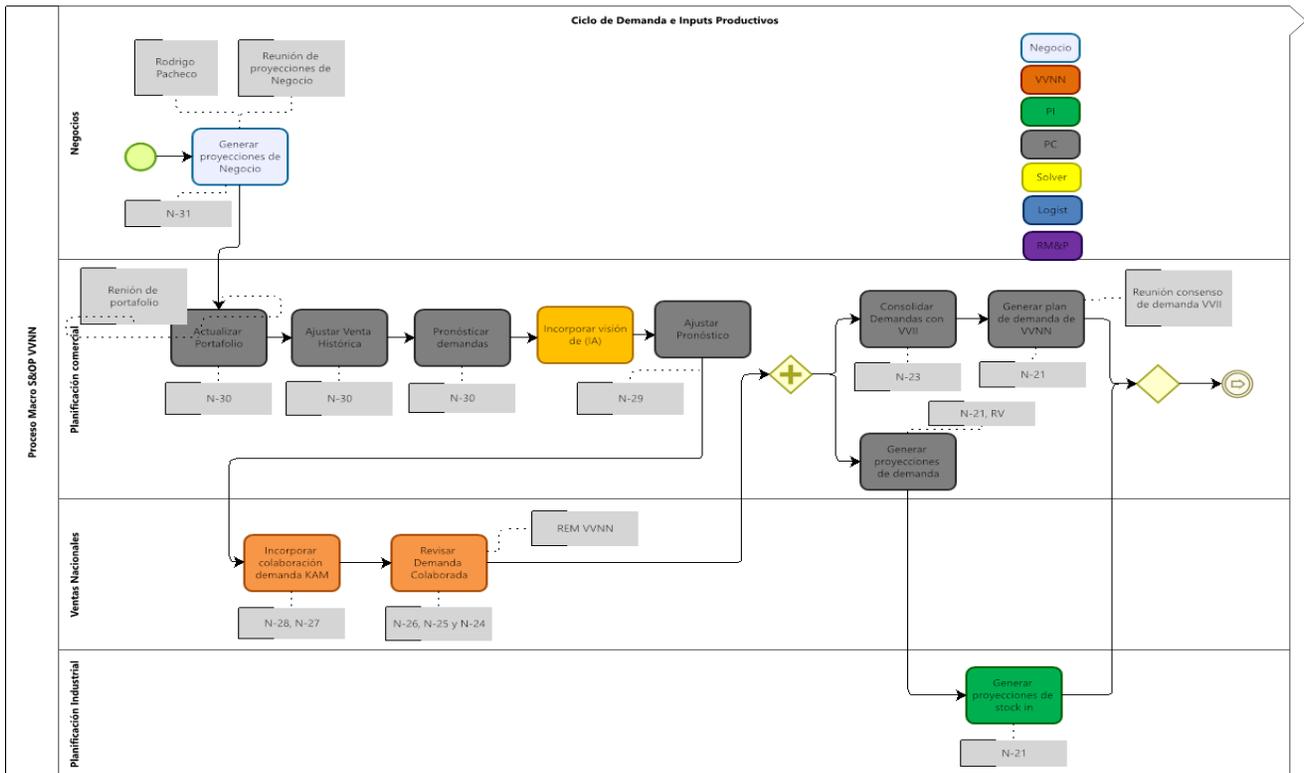
- Incluir la mirada de otros equipos para consolidar el proceso.
- Implementar acciones para lograr el escenario deseable.
- Agentes de cambio que promuevan la importancia de estas mejoras y logren movilizar a los equipos para lograr implementar lo propuesto en tiempo y forma, controlando los hitos y plazos establecidos.
- Recursos claves, horas de trabajo operativo, compromiso, promover el espíritu de mejora.
- Extender el diagnóstico de demanda a otras etapas del ciclo.

Con el proceso rediseñado se espera que la dinámica general se comporte de la siguiente manera:



Figura 28. Proceso de estimación de demanda propuesto. Elaboración propia.

Y al modelar este rediseño en BPMN, se puede apreciar la dinámica en la siguiente figura:



Powered by
bizagi
Modeler

Figura 29. Proceso de estimación de demanda propuesto en notación BPMN. Elaboración propia.

Se puede apreciar que sólo se incorporó una actividad que trata de incorporar la visión del modelo de red neuronal artificial o de la Inteligencia Artificial (IA), en dicha instancia se espera que luego de construir los pronósticos estadísticos, y haber sido manipulados e intencionados por el planificador comercial agregando su cuota de experiencia del negocio, se analicen las toneladas que muestra el algoritmo de RNA y se ajusten las que se tienen en ese momento en caso de ser necesario y que al planificador le haga sentido, vale decir que esto se propone para los principales materiales de cada sector que han sido expuestos en este trabajo (a nivel de agregación 2) , en estricto rigor, sería una colaboración más para construir el plan consensuado de demanda de cada mes.

9.2. Tecnologías habilitantes

Acorde a los objetivos del proyecto y a la dirección de cambio que se propone, especialmente respecto a la variable de apoyo computacional, se decide que la tecnología habilitante para el rediseño es la mezcla entre el software utilizado actualmente SAP IBP y tecnologías y entornos de programación dados por la nube, en específico Google Colab. Como se comentó previamente la solución SAP Integrated Business Planning (SAP IBP) es una plataforma cloud basada en tecnología “in memory” de SAP HANA y diseñada para integrar los procesos de planificación avanzada de la

cadena de suministro. Combina la potencia de SAP HANA, que permite alojar gran volumen de información y rapidez en los cálculos, con Excel como interface de usuario. Presenta en resumen las siguientes funcionalidades:

GESTIÓN DE LA DEMANDA

Modelos estadísticos y algoritmos de Machine Learning para obtener una transparencia total de la demanda previsiones a corto, medio y largo plazo.

OPTIMIZACIÓN DEL INVENTARIO

Establece objetivos de inventario óptimos que te permitan maximizar los beneficios y, al mismo tiempo, dejar un buffer que te ayude a satisfacer una demanda inesperada...

PLANIFICACIÓN DE VENTAS Y OPERACIONES

Consigue un plan de ventas y operaciones interdepartamental que equilibre el inventario, los niveles de servicio y la rentabilidad.

PLANIFICACIÓN DE RESPONSE AND SUPPLY

Optimiza la eficiencia de los recursos mediante la creación de planes de suministro basados en las demandas prioritarias y las restricciones de la cadena de suministro.

DEMAND DRIVEN REPLENISHMENT

Controla la variabilidad de la cadena de suministro y mejora el flujo de materiales utilizando las últimas técnicas en planificación

Entre las características se comenta que:

- Solución cloud (SaaS)
- Visibilidad total de extremo a extremo de toda su cadena de suministro
- Funcionalidades de planificación avanzada que cubren todos los aspectos de las cadenas de suministro más complejas
- Experiencia de usuario moderna y sencilla basada en Microsoft Excel y SAP Fiori
- Rápida implementación
- Personalizable

En las figuras 30 y 31, se muestra el interfaz que tiene un planificador comercial y los distintos equipos de la cadena de suministro de Agrosuper.

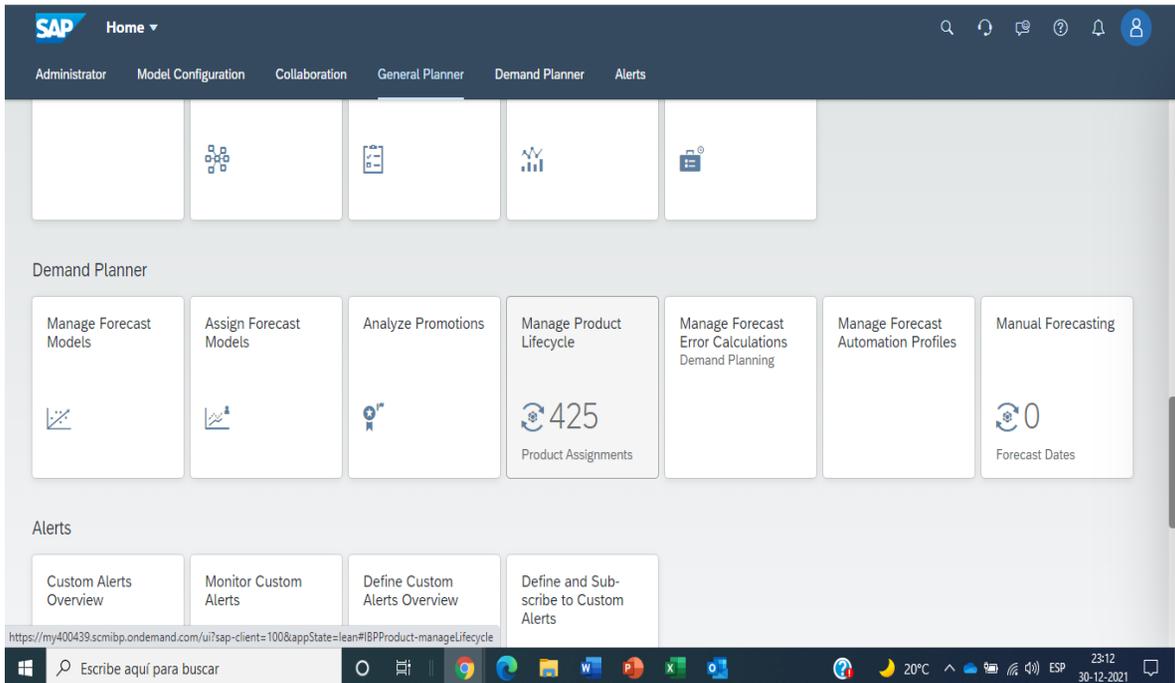


Figura 30. Interfaz de usuario dada por SAP Fiori.

	E	F	G	H	I	J	K	L	M
				W01 2021	W02 2021	W03 2021	W03 2021	W03 2021	W04
	Producto grupo	Nivel 2	Escenario	Demanda Histórica Kg	Plan Operativo	Demanda Histórica Kg	Plan Operativo	Demanda Histórica Kg	Plan Operativo
2	Cecina	Arrollado	Prueba pronóstico operativo	7634,67	7023,630563	7744,41	6579,829497	6871,03	6446,244842
3	Cecina	Fiambre	Prueba pronóstico operativo	45242,69	46140,68426	40108,01	43407,68317	40568,21	43119,50801
4	Cecina	Jamon	Prueba pronóstico operativo	182504,7	136588,6338	177141,21	130749,4226	149503,86	130241,3507
5	Cecina	Mortadela	Prueba pronóstico operativo	46760,6	48396,91992	45243,08	45415,38154	41193,2	45062,8974
6	Cecina	Parrilleros	Prueba pronóstico operativo	93793,32	48208,77389	77691,28	47427,7985	62835,89	46741,82357
7	Cecina	Pate	Prueba pronóstico operativo	1472	1086,015377	1634	1146,187305	1192	1143,033683
8	Cecina	Pechuga	Prueba pronóstico operativo	213301,37	138606,7564	206338,78	133488,1271	204328,98	133667,4751
9	Cecina	Prod. Espec.	Prueba pronóstico operativo	312,6	0	947,62	0	994,45	0
10	Cecina	Salame	Prueba pronóstico operativo	3624,56	863,158234	3563,76	869,220617	4174,44	865,886208
11	Cecina	Salchicha	Prueba pronóstico operativo	31845,31	25045,03512	33870	24294,30463	26459,2	24224,28919
12	Cerdo	Cabeza	Prueba pronóstico operativo				0,00033		0,000347
13	Cerdo	Chuleta	Prueba pronóstico operativo	436525,67	453409,5207	466110,68	430221,5772	498888,04	452539,2084
14	Cerdo	Cost-Pec	Prueba pronóstico operativo	262541,03	192237,5496	235431,58	216711,9773	218535,5	225955,336
15	Cerdo	Cueros	Prueba pronóstico operativo	10806,87	34548,54971	26931,32	28686,79529	29625,21	28686,79554
16	Cerdo	Filete	Prueba pronóstico operativo	10350,27	7373,313105	8552,83	8835,615756	8852,55	8803,60889
17	Cerdo	Grasa	Prueba pronóstico operativo	92847,05	124304,4292	135965,23	117766,8854	126216,62	117766,8854
18	Cerdo	Grasas	Prueba pronóstico operativo		14306,65366		0,001218		595
19	Cerdo	Huesos	Prueba pronóstico operativo	37313,95	29978,02463	42655,21	32199,86696	40731,39	32199,87449

Figura 31. Interfaz de usuario en Microsoft Excel dada por SAP IBP.

Los beneficios de implementación han sido notorios y se espera que con el rediseño siga aportando en estos, a continuación, se nombran algunos de estos:

- Reduce los niveles de inventario

- Disminuye los costes de logísticos, de producción y transporte
- Menores tiempos de entrega
- Reduce el ciclo de conversión de efectivo
- Incrementa el nivel de servicio
- Incrementa la rotación
- Mejorar la visibilidad de toda la cadena de suministro
- Racionaliza el SKU's y el modelo logístico
- Tiempos de respuesta muy rápidos en entornos con mucho volumen de datos

A continuación, se presenta el espectro que contiene los perfiles y modelos estadísticos provistos por SAP IBP. Tal como indica la figura, los modelos enumerados en cada cuadro representan los que pueden tratar ese patrón en los datos (por ejemplo la tendencia, la estacionalidad, la intermitencia y la insuficiencia de datos)

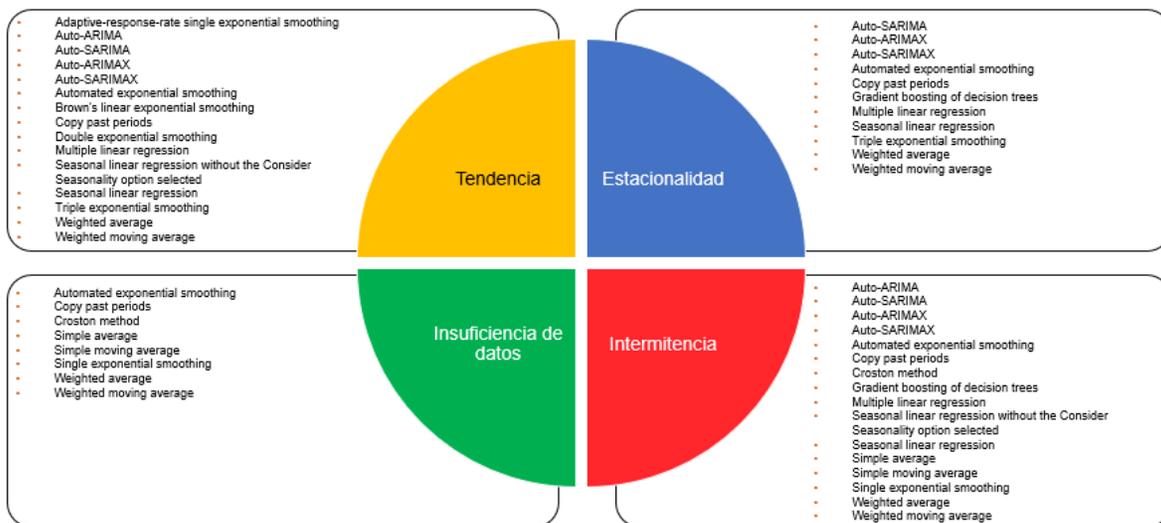


Figura 32. Perfiles y modelos estadísticos provistos por SAP IBP.

Por otro lado, se usará las plataformas open source de Google, específicamente Google Colaboratory para la programación, ejecución y desarrollo del modelo de red neuronal artificial, pues la herramienta de SAP IBP no permite la integración de este nuevo modelo en su interface. Sería utilizado de manera complementaria a los procesos descritos previamente.

Colaboratory, o "Colab" para abreviar, es un producto de Google Research. Permite a cualquier usuario escribir y ejecutar código arbitrario de Python en el navegador. Es especialmente adecuado para tareas de aprendizaje automático, análisis de datos y educación.

Colab nos permite conectarnos a un servidor de Jupyter que tengamos configurado en local si vamos al botón de «Conectar» y «Conectar a un entorno de ejecución local». Esto permite acceder a el sistema de archivos local, con las versiones que se quieran y que se tengan instaladas de manera local.

Dentro del mundo del Data Science, existen iniciativas muy interesantes, y una de las que más no puede interesar, además de todas las opciones formativas y herramientas disponibles, son los Google Colab.

Colab es un servicio cloud, basado en los Notebooks de Jupyter, que permite el uso gratuito de las GPUs y TPUs de Google, con librerías como: Scikit-learn, PyTorch, TensorFlow, Keras y OpenCV. Todo ello con bajo Python 2.7 y 3.6, que aún no está disponible para R y Scala.

Aunque tiene algunas limitaciones, que pueden consultarse en su página de FAQ, es una herramienta ideal, no solo para practicar y mejorar nuestros conocimientos en técnicas y herramientas de Data Science, sino también para el desarrollo de aplicaciones (pilotos) de machine learning y deep learning, sin tener que invertir en recursos hardware o del Cloud.

Con Colab se pueden crear notebooks o importar los que ya tengamos creados, además de compartirlos y exportarlos cuando queramos. Esta fluidez a la hora de manejar la información también es aplicable a las fuentes de datos que usemos en nuestros proyectos (notebooks), de modo que podremos trabajar con información contenida en nuestro propio Google Drive, unidad de almacenamiento local, github e incluso en otros sistemas de almacenamiento cloud, como el de Amazon.

10. Resultados de Plan Piloto

Luego del diseño y aplicación del modelo de red neuronal desarrollado previamente, se estructuró un plan “piloto” de parte de la propuesta de mejora que se plantea, específicamente con el objetivo de medir y comparar la efectividad de las proyecciones del modelo de red neuronal artificial desarrollado y el set de modelos estadísticos tradicionales – los más utilizados por el área de S&OP, provistos por SAP IBP –. En el plan, se evaluará si los MAPE de predicción de demanda del modelo propuesto son más efectivos y se usará para calcular beneficios potenciales y hacer la evaluación económica del proyecto en los próximos capítulos bajo ciertos supuestos que serán asumidos para esto.

Para esto, con una misma base muestral se hicieron pronósticos estadísticos “puros” para las siguientes semanas del año 2021 (5 semanas de análisis, desde la semana 43 a la 47 del presente año, 2021). Los modelos analizados y utilizados fueron: Auto-ARIMA; Auto-SARIMA; alisamiento exponencial automático; regresión lineal temporal y el modelo de RNA. Estos pronósticos serán contrastados semana a semana en una reunión con el equipo de S&OP en donde se analizará los modelos que tienen una mejor efectividad a la hora de hacer predicciones y se tomarán decisiones respecto a las futuras predicciones y si el proyecto representa una oportunidad importante para el área y para la empresa en general.

Esta base muestral corresponde a series temporales de demanda semanal para distintas categorías de productos (nivel 2), de los sectores de cerdo, pollo, pavo, cecina y elaborado. En la siguiente tabla se muestra un ejemplo con las categorías de productos para el sector del cerdo junto con sus respectivas estimaciones. Cabe destacar que aún no se obtienen resultados dado que la demanda real para esta semana estará consolidada el día 1 de noviembre del presente año, y así sucesivamente cada lunes.

En el presente estudio se configuró la RNA desarrollada en el capítulo 5 mediante entornos ofrecidos por Google Colab en un lenguaje de programación de Python, mientras que los demás modelos estadísticos fueron obtenidos por la interfaz de Fiori de SAP IBP. Por un lado, los resultados de esta investigación permitirán a la industria contar con un mayor conjunto de elementos de juicio para la toma de decisión final en la gestión de sus inventarios, en la planificación misma. Esto permitirá una mejor gestión y conciliación entre la cantidad demandada y la cantidad producida. Por otro lado, es evidente todavía la dificultad en su aplicación directa con la automatización de la previsión mediante RNA en los actuales sistemas ERP (Enterprise Resource Planning) utilizados en Agrosuper. (SAP IBP)

A continuación se presentan las tablas con los resultados del desempeño medido a través del indicador de media del error absoluto en porcentaje (MAPE) de cada sector y de cada semana de estudio.

Semana 43

Sector	Material	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Cerdo	Chuleta	45%	4%	57%	2%	12%
Cerdo	Cost-Pec	35%	3%	59%	34%	14%
Cerdo	Cueros	50%	11%	59%	19%	19%
Cerdo	Filete	53%	6%	64%	17%	32%
Cerdo	Grasa	42%	3%	65%	6%	24%
Cerdo	Huesos	61%	9%	66%	23%	44%
Cerdo	Lomo	45%	11%	28%	10%	1%
Cerdo	Paleta	43%	33%	60%	17%	10%
Cerdo	Pernil	62%	15%	65%	50%	16%
Cerdo	Pierna	57%	4%	64%	16%	26%
Cerdo	Plancha	24%	64%	8%	90%	16%
Cerdo	Prolijado	32%	26%	52%	24%	3%
Cerdo	Recortes	27%	6%	46%	7%	4%
Cerdo	Reprod	50%	7%	60%	1%	76%
Cerdo	Subprod	28%	31%	39%	25%	8%

Tabla 14. Resultados de los MAPE de predicciones por modelo de los materiales del sector cerdo en la semana 43 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Pollo	Ala	58%	7%	67%	0%	28%
Pollo	Cazuela	64%	21%	69%	19%	528%
Pollo	Con Menudencia	55%	3%	72%	4%	1%
Pollo	Filete	45%	0%	41%	4%	7%
Pollo	Menudencias	68%	6%	72%	20%	14%
Pollo	Patas	76%	16%	82%	49%	63%
Pollo	Pechuga	53%	2%	63%	15%	14%
Pollo	Pechuga Desh	63%	27%	74%	17%	23%
Pollo	Reproductor	60%	30%	54%	3%	158%
Pollo	Trutro	48%	8%	64%	18%	13%
Pollo	Trutro Deshuesado	78%	36%	74%	61%	9%

Tabla 15. Resultados de los MAPE de predicciones por modelo de los materiales del sector pollo en la semana 43 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Pavo	Ala	66%	6%	70%	7%	36%
Pavo	Menudencias	38%	13%	56%	6%	9%
Pavo	Pech Desh	53%	12%	61%	17%	25%
Pavo	Trutro	62%	4%	68%	0%	23%
Pavo	Trutro Desh	62%	7%	68%	6%	23%

Tabla 16. Resultados de los MAPE de predicciones por modelo de los materiales del sector pavo en la semana 43 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Cecina	Arrollado	51%	4%	52%	4%	37%
Cecina	Fiambre	43%	2%	50%	3%	4%
Cecina	Jamon	52%	15%	59%	21%	22%
Cecina	Mortadela	48%	13%	54%	9%	23%
Cecina	Parrilleros	10%	11%	21%	8%	4%
Cecina	Pate	39%	34%	54%	31%	30%
Cecina	Pechuga	50%	8%	57%	14%	31%
Cecina	Prod. Espec.	61%	13%	65%	27%	17%
Cecina	Salame	38%	32%	61%	40%	30%
Cecina	Salchicha	39%	5%	41%	28%	107%

Tabla 17. Resultados de los MAPE de predicciones por modelo de los materiales del sector cecina en la semana 43 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Elaborado	Empanadas	30%	4%	34%	26%	52%
Elaborado	Empanizado	34%	9%	46%	7%	3%
Elaborado	Hamburguesa	49%	22%	58%	23%	33%
Elaborado	Moldeado	44%	11%	48%	16%	20%
Elaborado	Molido	51%	14%	58%	36%	41%
Elaborado	Pizza	48%	26%	59%	13%	56%

Tabla 18. Resultados de los MAPE de predicciones por modelo de los materiales del sector elaborado en la semana 43 del año 2021.

Semana 44

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Cerdo	Chuleta	38%	10%	56%	10%	6%
Cerdo	Cost-Pec	22%	14%	55%	14%	11%
Cerdo	Cueros	21%	35%	19%	34%	8%
Cerdo	Filete	51%	37%	57%	10%	39%
Cerdo	Grasa	24%	23%	49%	36%	6%
Cerdo	Huesos	52%	4%	62%	6%	39%
Cerdo	Lomo	31%	26%	45%	30%	7%
Cerdo	Paleta	35%	44%	57%	6%	9%
Cerdo	Pernil	58%	18%	66%	51%	7%
Cerdo	Pierna	51%	3%	59%	10%	29%
Cerdo	Plancha	128%	202%	42%	257%	90%
Cerdo	Prolijado	23%	30%	50%	27%	4%
Cerdo	Recortes	1%	25%	33%	27%	25%
Cerdo	Reprod	39%	3%	49%	21%	103%
Cerdo	Subprod	23%	29%	45%	33%	1%

Tabla 19. Resultados de los MAPE de predicciones por modelo de los materiales del sector cerdo en la semana 44 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Pollo	Ala	57%	9%	69%	28%	5%
Pollo	Cazuela	70%	36%	75%	13%	387%
Pollo	Con Menudencia	52%	5%	74%	8%	3%
Pollo	Filete	33%	13%	53%	1%	0%
Pollo	Menudencias	64%	1%	70%	4%	8%
Pollo	Patas	70%	4%	78%	13%	59%
Pollo	Pechuga	52%	9%	64%	22%	0%
Pollo	Pechuga Desh	49%	2%	66%	6%	2%
Pollo	Reproductor	13%	23%	0%	2%	313%
Pollo	Trutro	42%	8%	63%	26%	9%
Pollo	Trutro Deshuesado	78%	38%	77%	54%	12%

Tabla 20. Resultados de los MAPE de predicciones por modelo de los materiales del sector pollo en la semana 44 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Pavo	Ala	59%	5%	66%	6%	27%
Pavo	Menudencias	9%	58%	39%	45%	3%
Pavo	Pech Desh	50%	23%	60%	14%	25%
Pavo	Trutro	57%	3%	65%	5%	22%
Pavo	Trutro Desh	56%	3%	64%	2%	17%

Tabla 22. Resultados de los MAPE de predicciones por modelo de los materiales del sector pavo en la semana 44 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Cecina	Arrollado	51%	3%	53%	6%	35%
Cecina	Fiambre	35%	17%	41%	9%	11%
Cecina	Jamon	40%	1%	51%	6%	6%
Cecina	Mortadela	34%	8%	43%	10%	63%
Cecina	Parrilleros	17%	6%	28%	4%	1%
Cecina	Pate	18%	89%	45%	72%	6%
Cecina	Pechuga	40%	8%	50%	1%	19%
Cecina	Prod. Espec.	58%	7%	64%	12%	19%
Cecina	Salame	14%	13%	57%	79%	9%
Cecina	Salchicha	45%	6%	46%	31%	51%

Tabla 23. Resultados de los MAPE de predicciones por modelo de los materiales del sector cecina en la semana 44 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Elaborado	Empanadas	29%	4%	36%	22%	53%
Elaborado	Empanizado	22%	4%	39%	7%	4%
Elaborado	Hamburguesa	29%	6%	42%	6%	10%
Elaborado	Moldeado	34%	1%	46%	3%	10%
Elaborado	Molido	49%	9%	55%	31%	46%
Elaborado	Pizza	34%	20%	29%	5%	50%

Tabla 24. Resultados de los MAPE de predicciones por modelo de los materiales del sector elaborado en la semana 44 del año 2021.

Semana 45

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Cerdo	Chuleta	41%	0%	59%	0%	9%
Cerdo	Cost-Pec	25%	8%	59%	6%	12%
Cerdo	Cueros	4%	59%	13%	47%	26%
Cerdo	Filete	61%	20%	70%	10%	51%
Cerdo	Grasa	18%	27%	48%	33%	2%
Cerdo	Huesos	57%	8%	66%	23%	46%
Cerdo	Lomo	47%	7%	58%	16%	8%
Cerdo	Paleta	40%	28%	61%	20%	19%
Cerdo	Pernil	51%	11%	62%	48%	6%
Cerdo	Pierna	53%	4%	63%	13%	39%
Cerdo	Plancha	311%	444%	139%	593%	254%
Cerdo	Prolijado	42%	6%	66%	8%	16%
Cerdo	Recortes	6%	14%	40%	24%	19%
Cerdo	Reprod	38%	10%	55%	3%	68%
Cerdo	Subprod	2%	63%	34%	57%	20%

Tabla 25. Resultados de los MAPE de predicciones por modelo de los materiales del sector cerdo en la semana 45 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Pollo	Ala	62%	19%	74%	21%	0%
Pollo	Cazuela	72%	42%	84%	37%	255%
Pollo	Con Menudencia	50%	8%	76%	12%	5%
Pollo	Filete	49%	17%	66%	25%	17%
Pollo	Menudencias	71%	20%	77%	29%	22%
Pollo	Patas	80%	34%	80%	51%	78%
Pollo	Pechuga	54%	15%	68%	27%	2%
Pollo	Pechuga Desh	59%	22%	75%	17%	23%
Pollo	Reproductor	70%	63%	63%	47%	19%
Pollo	Trutro	50%	23%	70%	31%	30%
Pollo	Trutro Deshuesado	78%	39%	78%	67%	7%

Tabla 26. Resultados de los MAPE de predicciones por modelo de los materiales del sector pollo en la semana 45 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Pavo	Ala	61%	5%	70%	10%	37%
Pavo	Menudencias	31%	17%	54%	16%	8%
Pavo	Pech Desh	53%	28%	62%	15%	38%
Pavo	Trutro	61%	12%	71%	8%	32%
Pavo	Trutro Desh	58%	5%	67%	3%	24%

Tabla 27. Resultados de los MAPE de predicciones por modelo de los materiales del sector pavo en la semana 45 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Cecina	Arrollado	52%	8%	56%	1%	36%
Cecina	Fiambre	50%	15%	59%	19%	12%
Cecina	Jamon	47%	13%	58%	20%	22%
Cecina	Mortadela	47%	15%	54%	7%	29%
Cecina	Parrilleros	21%	17%	38%	15%	12%
Cecina	Pate	19%	76%	45%	57%	10%
Cecina	Pechuga	45%	3%	55%	9%	29%
Cecina	Prod. Espec.	57%	3%	59%	25%	15%
Cecina	Salame	16%	18%	57%	61%	25%
Cecina	Salchicha	48%	12%	49%	29%	48%

Tabla 28. Resultados de los MAPE de predicciones por modelo de los materiales del sector cecina en la semana 45 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Elaborado	Empanadas	36%	15%	42%	2%	55%
Elaborado	Empanizado	38%	18%	52%	18%	9%
Elaborado	Hamburguesa	41%	14%	54%	16%	24%
Elaborado	Moldeado	45%	18%	57%	21%	29%
Elaborado	Molido	65%	38%	69%	53%	3%
Elaborado	Pizza	52%	32%	66%	23%	61%

Tabla 29. Resultados de los MAPE de predicciones por modelo de los materiales del sector elaborado en la semana 45 del año 2021.

Semana 46

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Cerdo	Chuleta	39%	2%	57%	3%	13%
Cerdo	Cost-Pec	14%	24%	53%	23%	4%
Cerdo	Cueros	14%	39%	24%	31%	2%
Cerdo	Filete	45%	40%	58%	6%	29%
Cerdo	Grasa	30%	5%	63%	13%	15%
Cerdo	Huesos	54%	4%	63%	16%	41%
Cerdo	Lomo	41%	2%	54%	13%	4%
Cerdo	Paleta	35%	30%	60%	15%	11%
Cerdo	Pernil	54%	18%	66%	50%	7%
Cerdo	Pierna	50%	1%	61%	13%	37%
Cerdo	Plancha	176%	264%	71%	327%	128%
Cerdo	Prolijado	29%	14%	57%	8%	1%
Cerdo	Recortes	2%	15%	38%	23%	20%
Cerdo	Reprod	38%	2%	55%	14%	95%
Cerdo	Subprod	20%	96%	22%	78%	39%

Tabla 30. Resultados de los MAPE de predicciones por modelo de los materiales del sector cerdo en la semana 46 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Pollo	Ala	55%	11%	70%	29%	13%
Pollo	Cazuela	70%	41%	86%	24%	233%
Pollo	Con Menudencia	49%	11%	76%	15%	10%
Pollo	Filete	40%	4%	60%	14%	1%
Pollo	Menudencias	65%	11%	74%	13%	12%
Pollo	Patas	75%	21%	80%	45%	77%
Pollo	Pechuga	60%	28%	72%	24%	14%
Pollo	Pechuga Desh	58%	30%	75%	23%	27%
Pollo	Reproductor	76%	73%	64%	65%	37%
Pollo	Trutro	48%	21%	69%	7%	24%
Pollo	Trutro Deshuesado	73%	27%	69%	35%	0%

Tabla 31. Resultados de los MAPE de predicciones por modelo de los materiales del sector pollo en la semana 46 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Pavo	Ala	61%	5%	69%	5%	36%
Pavo	Menudencias	29%	18%	55%	2%	9%
Pavo	Pech Desh	54%	19%	59%	24%	39%
Pavo	Trutro	54%	2%	65%	1%	19%
Pavo	Trutro Desh	61%	15%	71%	14%	28%

Tabla 32. Resultados de los MAPE de predicciones por modelo de los materiales del sector pavo en la semana 46 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Cecina	Arrollado	50%	7%	57%	3%	36%
Cecina	Fiambre	45%	7%	56%	14%	1%
Cecina	Jamon	47%	15%	59%	21%	19%
Cecina	Mortadela	41%	6%	50%	3%	51%
Cecina	Parrilleros	11%	9%	28%	0%	6%
Cecina	Pate	41%	27%	59%	15%	33%
Cecina	Pechuga	44%	3%	56%	9%	29%
Cecina	Prod. Espec.	71%	36%	75%	41%	7%
Cecina	Salame	3%	8%	53%	80%	13%
Cecina	Salchicha	41%	0%	42%	28%	76%

Tabla 33. Resultados de los MAPE de predicciones por modelo de los materiales del sector cecina en la semana 46 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Elaborado	Empanadas	39%	19%	48%	3%	59%
Elaborado	Empanizado	30%	9%	45%	9%	2%
Elaborado	Hamburguesa	38%	11%	40%	12%	19%
Elaborado	Moldeado	46%	19%	58%	24%	28%
Elaborado	Molido	61%	31%	65%	50%	20%
Elaborado	Pizza	56%	31%	55%	37%	67%

Tabla 34. Resultados de los MAPE de predicciones por modelo de los materiales del sector elaborado en la semana 46 del año 2021.

Semana 47

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Cerdo	Chuleta	37%	4%	57%	6%	13%
Cerdo	Cost-Pec	5%	35%	50%	30%	4%
Cerdo	Cueros	29%	13%	46%	23%	6%
Cerdo	Filete	52%	4%	67%	19%	39%
Cerdo	Grasa	3%	53%	45%	68%	20%
Cerdo	Huesos	55%	8%	66%	18%	43%
Cerdo	Lomo	44%	7%	60%	12%	10%
Cerdo	Paleta	36%	27%	62%	18%	10%
Cerdo	Pernil	53%	17%	66%	52%	4%
Cerdo	Pierna	50%	3%	59%	14%	38%
Cerdo	Plancha	468%	648%	253%	715%	385%
Cerdo	Prolijado	25%	18%	56%	15%	3%
Cerdo	Recortes	14%	0%	47%	15%	3%
Cerdo	Reprod	46%	11%	66%	2%	56%
Cerdo	Subprod	4%	56%	35%	61%	18%

Tabla 35. Resultados de los MAPE de predicciones por modelo de los materiales del sector cerdo en la semana 47 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Pollo	Ala	49%	7%	67%	13%	18%
Pollo	Cazuela	67%	35%	80%	34%	163%
Pollo	Con Menudencia	45%	8%	73%	11%	3%
Pollo	Filete	43%	10%	63%	14%	9%
Pollo	Menudencias	64%	10%	74%	18%	12%
Pollo	Patas	76%	27%	84%	54%	81%
Pollo	Pechuga	53%	18%	69%	29%	6%
Pollo	Pechuga Desh	59%	27%	75%	20%	30%
Pollo	Reproductor	3%	6%	23%	105%	324%
Pollo	Trutro	42%	14%	66%	23%	21%
Pollo	Trutro Deshuesado	71%	24%	71%	55%	7%

Tabla 36. Resultados de los MAPE de predicciones por modelo de los materiales del sector pollo en la semana 47 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Pavo	Ala	56%	4%	68%	3%	29%
Pavo	Menudencias	30%	14%	55%	1%	0%
Pavo	Pech Desh	68%	45%	68%	49%	57%
Pavo	Trutro	61%	14%	72%	8%	31%
Pavo	Trutro Desh	56%	5%	67%	5%	17%

Tabla 37. Resultados de los MAPE de predicciones por modelo de los materiales del sector pavo en la semana 47 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Cecina	Arrollado	45%	3%	53%	15%	31%
Cecina	Fiambre	44%	7%	55%	7%	1%
Cecina	Jamon	43%	9%	57%	15%	15%
Cecina	Mortadela	43%	10%	52%	7%	42%
Cecina	Parrilleros	9%	11%	10%	19%	4%
Cecina	Pate	12%	85%	50%	67%	4%
Cecina	Pechuga	38%	6%	51%	1%	23%
Cecina	Prod. Espec.	57%	4%	66%	12%	17%
Cecina	Salame	7%	13%	57%	84%	24%
Cecina	Salchicha	47%	11%	52%	38%	52%

Tabla 38. Resultados de los MAPE de predicciones por modelo de los materiales del sector cecina en la semana 47 del año 2021.

Sector	Nivel 2	MAPE ARIMA	MAPE SARIMA	MAPE AEA	MAPE RLT	MAPE RNA
Elaborado	Empanadas	22%	5%	28%	29%	51%
Elaborado	Empanizado	20%	3%	39%	9%	3%
Elaborado	Hamburguesa	26%	6%	44%	6%	6%
Elaborado	Moldeado	38%	8%	52%	10%	16%
Elaborado	Molido	59%	28%	64%	47%	17%
Elaborado	Pizza	23%	7%	48%	14%	46%

Tabla 39. Resultados de los MAPE de predicciones por modelo de los materiales del sector elaborado en la semana 47 del año 2021.

10.1. Análisis de resultados

Luego de analizar los resultados, se construye un ranking de los modelos predictivos el cual se puede observar en el gráfico siguiente, en donde se aprecia que el modelo que entrega una predicción más exacta durante las 5 semanas de estudio es el modelo SARIMA con un 40% de los casos, quedando el modelo de red neuronal en segundo lugar con un 31% de los casos con mejores predicciones, y el modelo de regresión lineal temporal con un 22% de los casos.

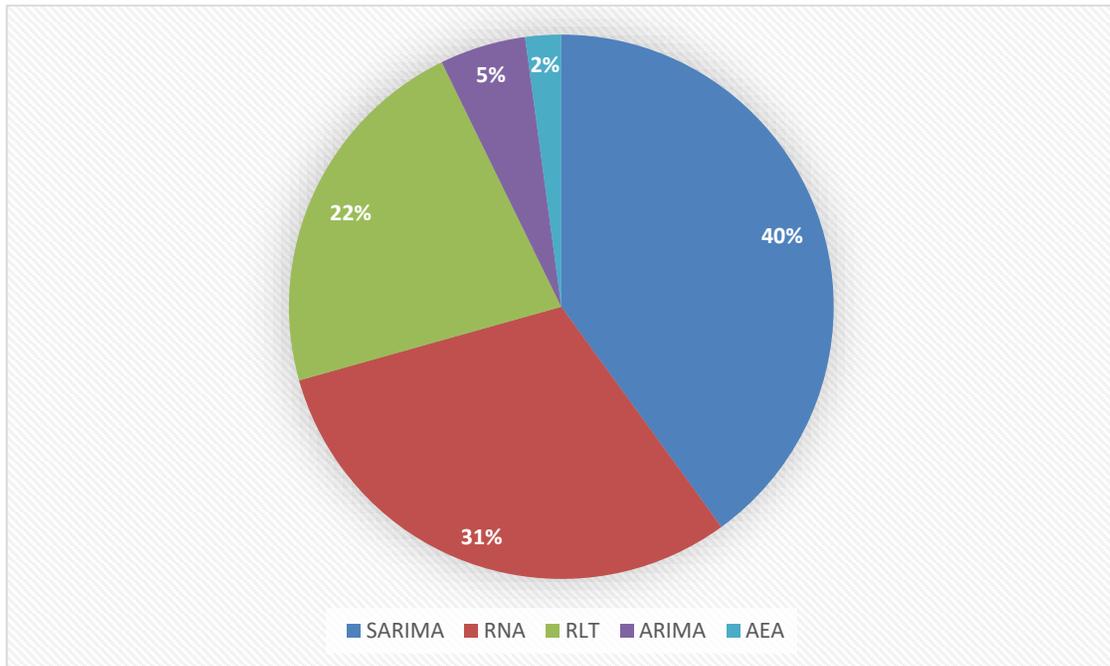


Gráfico 30. Ranking de efectividad en relación con el mejor modelo para predecir en las 5 semanas de estudio del año 2021.

Es necesario realizar un análisis en base a la segmentación de cada sector y contrastar los resultados de estos, por lo que se presenta un resumen con los datos para cada negocio (cerdo, pollo, pavo, cecina y elaborados). Además, se presenta el resumen de los resultados para cada material en donde se comparan los MAPE de la red neuronal artificial, el mejor MAPE de los modelos provistos por SAP IBP (SARIMA, RLT, ARIMA y AEA) y como variable adicional se compara el MAPE de la colaboración que hace ventas en el proceso, analizando la efectividad de cada involucrado en la construcción del plan consensuado de demanda.

Cerdo

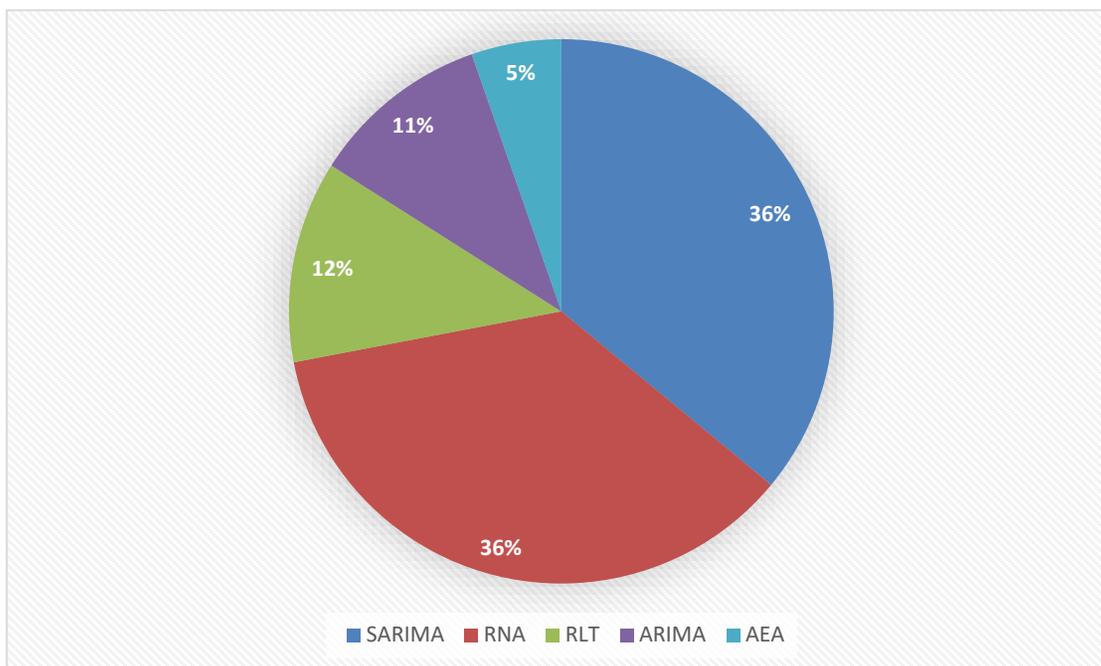


Gráfico 31. Ranking de efectividad en relación con el mejor modelo para predecir en las 5 semanas de estudio del año 2021 para el sector cerdo.

Se puede apreciar que el mejor modelo para este sector está disputado entre el modelo SARIMA y el modelo de red neuronal artificial con un 36% de los mejores pronósticos, sumando sólo un 28% entre los demás modelos (regresión lineal temporal, ARIMA y Alisamiento exponencial automático). Se puede notar en la tabla que 4 de los 15 materiales analizados de cerdo presentan un mejor MAPE a partir del modelo de red neuronal artificial, y en general los pronósticos de demanda con este modelo son buenos y competitivos respecto con los modelos ofrecidos por SAP IBP.

Nivel 2	MAPE IBP	MAPE CV	MAPE RNA
Chuleta	4%	7%	11%
Cost-pec	8%	10%	9%
Cueros	12%	170%	12%
Filete	7%	21%	38%
Grasa	10%	29%	13%
Huesos	7%	13%	43%
Lomo	10%	25%	6%
Paleta	15%	5%	12%
Pernil	16%	35%	8%
Pierna	3%	5%	34%
Plancha	106%	312%	175%
Prolijado	15%	19%	5%
Recortes	3%	48%	14%
Reprod	2%	11%	80%
Subprod	15%	82%	17%

Tabla 40. Resultados de los MAPE por material predicho del sector cerdo.

Pollo

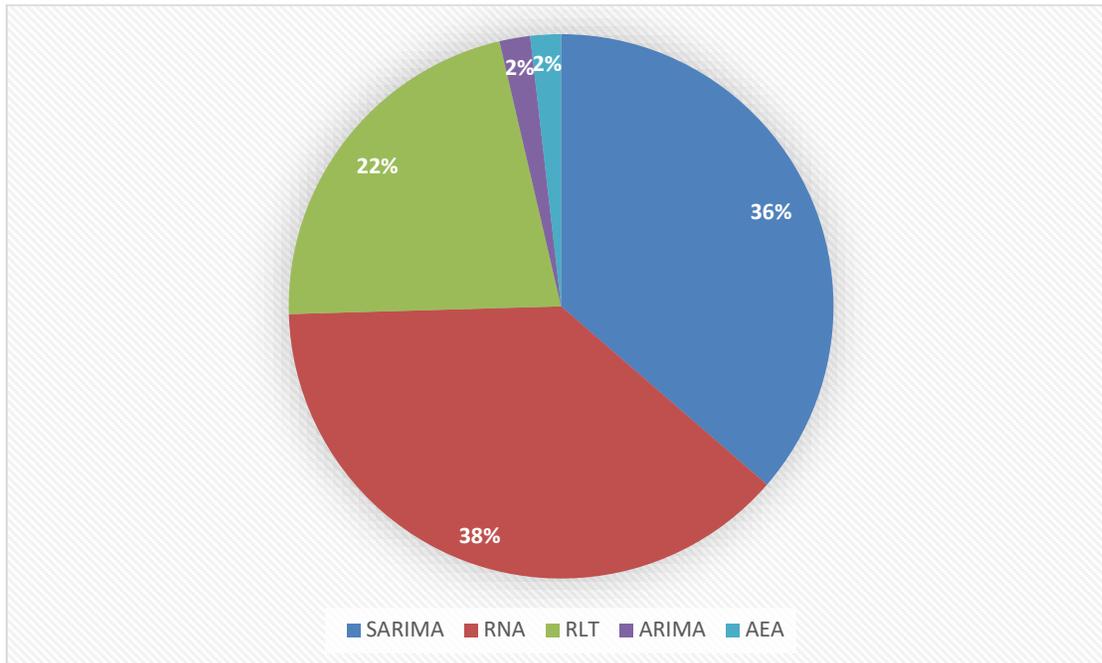


Gráfico 32. Ranking de efectividad en relación con el mejor modelo para predecir en las 5 semanas de estudio del año 2021 para el sector pollo.

Se puede notar en el gráfico que en el caso del pollo, la red neuronal artificial tuvo un mejor desempeño, situándose con la mayoría de las mejores predicciones de demanda con un 38% de los casos, en el que le sigue por muy cerca el modelo SARIMA con un 36% de los casos y el modelo de regresión lineal temporal con un 22% de los casos. Ahora bien, al analizar los distintos materiales se puede apreciar que 4 de los 11 materiales tuvieron un mejor desempeño con el modelo de RNA.

Nivel 2	MAPE IBP	MAPE CV	MAPE RNA
Ala	9%	10%	13%
Cazuela	25%	38%	303%
Con Menudencias	7%	10%	4%
Filete	7%	10%	7%
Menudencias	10%	11%	14%
Patas	20%	42%	73%
Pechuga	14%	11%	7%
Pechuga Desh	16%	10%	22%
Reproductor	23%	83%	170%
Trutro	12%	17%	19%
Trutro Desh	33%	37%	7%

Tabla 41. Resultados de los MAPE por material predicho del sector pollo.

Pavo

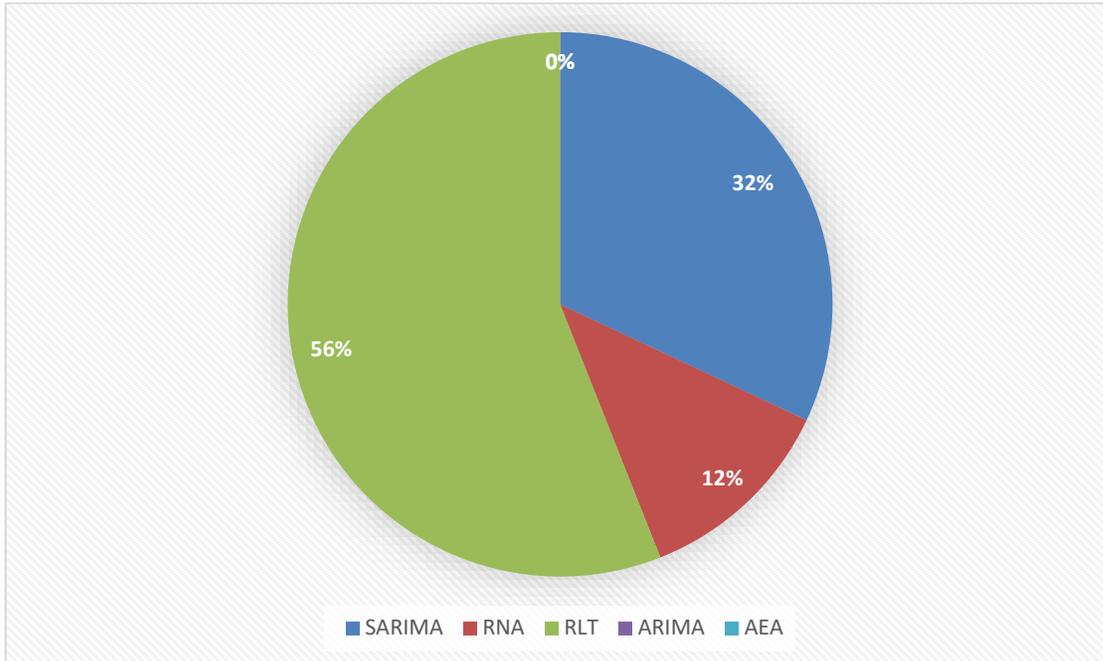


Gráfico 33. Ranking de efectividad en relación con el mejor modelo para predecir en las 5 semanas de estudio del año 2021 para el sector pavo.

Se puede notar en el gráfico que en el caso del pavo es bastante distinto a los anteriores, la red neuronal artificial tuvo un mal desempeño, situándose con el tercer lugar de las mejores predicciones de demanda con un 12% de los casos, el mejor desempeño para este sector lo tuvo el modelo de regresión lineal temporal con el 56% de los pronósticos realizados para este sector del negocio, mientras que el modelo SARIMA tuvo el segundo lugar con un 32% de los casos. Ahora bien, al analizar los distintos materiales es posible evidenciar que 1 de los 5 materiales de pavo tuvo un mejor desempeño aplicando el modelo de redes neuronales artificiales.

Nivel 2	MAPE IBP	MAPE CV	MAPE RNA
Ala	5%	4%	33%
Menudencias	7%	9%	6%
Pechuga Desh	21%	31%	37%
Trutro	4%	10%	25%
Trutro Desh	6%	8%	22%

Tabla 42. Resultados de los MAPE por material predicho del sector pavo.

Cecina

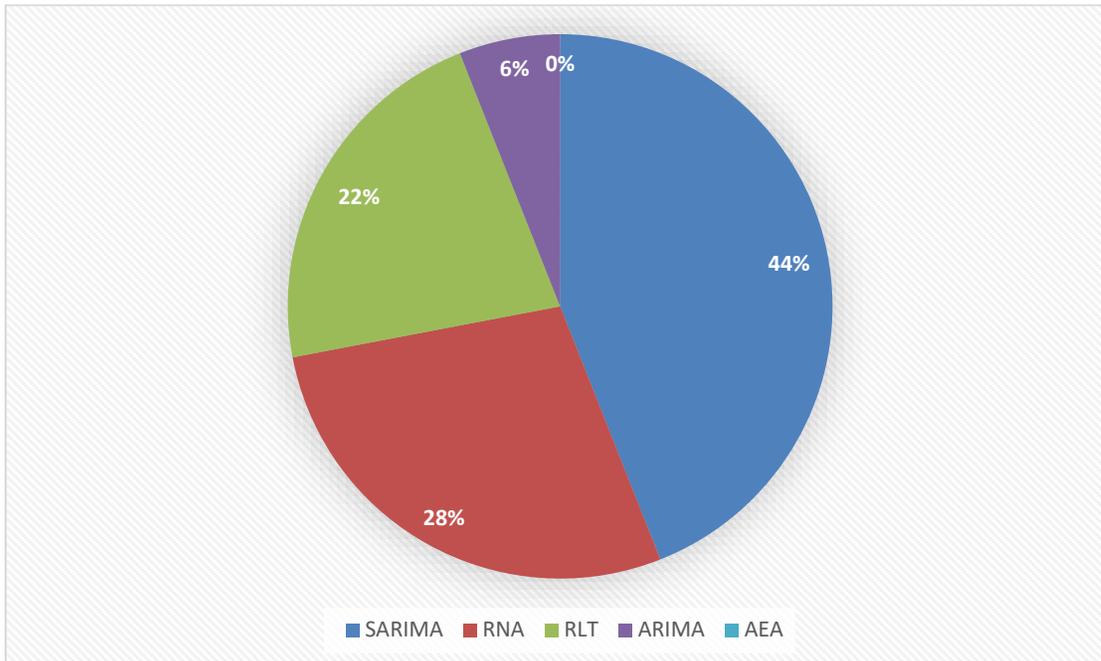


Gráfico 34. Ranking de efectividad en relación con el mejor modelo para predecir en las 5 semanas de estudio del año 2021 para el sector cecina.

Se aprecia en el gráfico que en el caso de las cecinas el mejor desempeño lo tuvo el modelo SARIMA con un 44% de los casos mejor predichos, la red neuronal artificial se sitúa en el segundo lugar en relación con los desempeños con un 28% de los casos, y el modelo de regresión lineal temporal se posiciona con el tercer lugar de las mejores predicciones de demanda con un 22%. Respecto al análisis por material, al analizarlos es posible ver que 3 de los 10 materiales tuvieron un mejor desempeño aplicando el modelo de redes neuronales artificiales.

Nivel 2	MAPE IBP	MAPE CV	MAPE RNA
Arrollado	3%	5%	35%
Fiambre	8%	7%	6%
Jamón	11%	16%	17%
Mortadela	7%	6%	42%
Parrilleros	8%	12%	5%
Pate	19%	22%	17%
Pechuga	3%	13%	26%
Prod. Espec.	13%	44%	15%
Salame	14%	25%	20%
Salchicha	7%	11%	67%

Tabla 43. Resultados de los MAPE por material predicho del sector cecina.

Elaborado

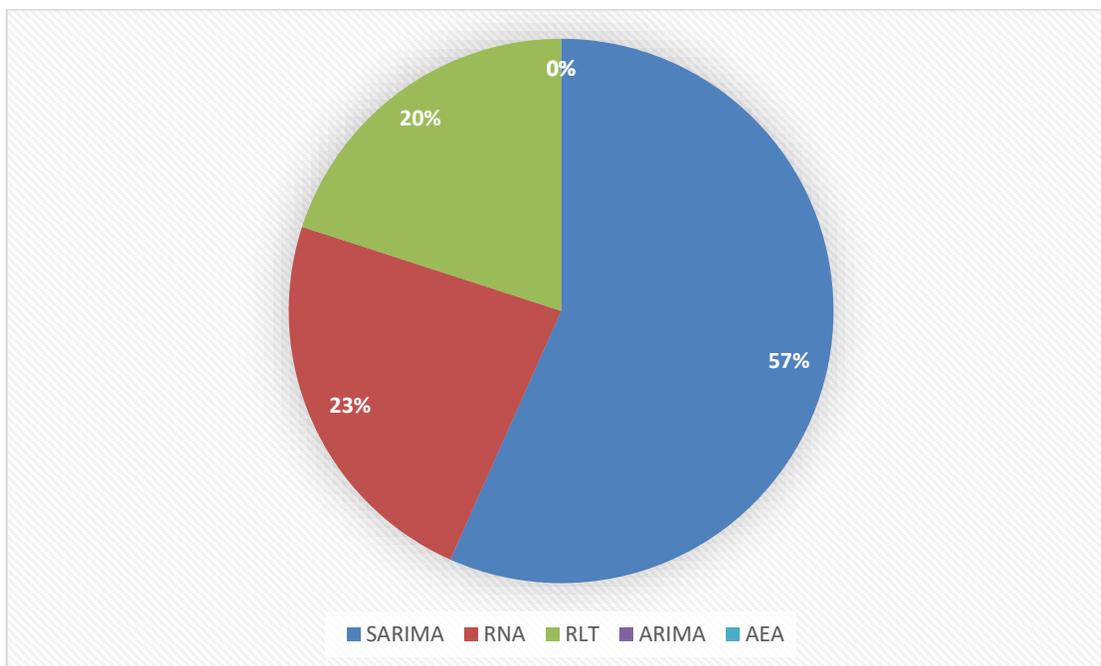


Gráfico 35. Ranking de efectividad en relación con el mejor modelo para predecir en las 5 semanas de estudio del año 2021 para el sector elaborado.

Se observa en el gráfico que en el caso del sector elaborados el mejor desempeño lo tuvo el modelo SARIMA con un 57% de los casos mejor predichos, la red neuronal artificial se sitúa en el segundo lugar con un desempeño del 23% de los casos con mejor pronóstico, y el modelo de regresión lineal temporal se posiciona con el tercer lugar con un 20%. Respecto al análisis por material, al analizarlos es posible ver que 1 de los 6 materiales tuvo un mejor desempeño aplicando el modelo de redes neuronales artificiales.

Nivel 2	MAPE IBP	MAPE CV	MAPE RNA
Empanadas	4%	12%	54%
Empanizado	8%	10%	4%
Hamburguesa	12%	13%	18%
Moldeado	11%	20%	21%
Molido	24%	7%	25%
Pizza	16%	25%	56%

Tabla 44. Resultados de los MAPE por material predicho del sector elaborado.

A juzgar por los resultados de este apartado, puede ser necesario que los algoritmos de aprendizaje automático aplicados a la previsión requieran investigación adicional para experimentar con ideas innovadoras y realizar ajustes para lograr predicciones más precisas.

11. Plan de Implementación

Para poner en funcionamiento el rediseño del proceso es necesario establecer un plan de implementación para coordinar las actividades necesarias con el fin de asegurar el éxito del desarrollo y de la implementación.

Propósito de la solución

El propósito de la solución es la implementación de una mejora al proceso de pronóstico de demanda realizado por la cadena de suministro, que contempla una serie de acciones claves que pueden implementarle además de la incorporación de un modelo analítico de aprendizaje automático para predecir la demanda de productos en base a la información histórica de los productos de Agrosuper, específicamente en los materiales y negocios de cerdo y pollo en los cuales tuvo un mejor desempeño al predecir.

Se concibe el modelo como una herramienta de colaboración complementaria a las actuales técnicas aplicadas en el proceso. Con esta solución se pretende entregar al planificador comercial una herramienta de apoyo a la toma de decisiones que le servirá como complemento para determinar el plan comercial final, ajustado con las visiones de los demás actores involucrados en el proceso.

Planificación de actividades

A continuación, se presenta las actividades y propuestas de mejora en formato de carta Gantt, con el fin de dar cumplimiento etapa por etapa, entregando los plazos establecidos, cabe destacar que en esta instancia del proceso se hace crucial incorporar el rol de la líder de S&OP, pues debe ser quien gestione estos cambios y que las cosas se cumplan tanto en tiempo como en calidad y forma, se requiere de una alta capacidad de liderazgo para enfrentar el desafío:

ETAPA: DEMANDA	ACTIVIDAD	C	ACCIONES	Responsable	Mes n	Mes n+1	Mes n+2	Mes n+3	Mes n+4	Mes n+5	Mes n+6	Mes n+7	Mes n+i		
I. Proyecciones del negocios	Infomación y lineamientos	B	Diseñar y comunicar formato estandarizado de presentación.	Negocios											
			Proyecciones acorde a la rentabilidad esperada proximos meses	Negocios											
			Comunicar situación futura de materiales foco (tendencias).	Negocios											
			Indicar holguras productivas, tendencias de precios y Mape/Fillrate de productos críticos para intensionar el proceso de Construcción de demanda.	Negocios											
			Enfasis nuevos lanzamientos.	Negocios											
			Aperturar lineamientos/estrategia para los diterentes canales/mercados.	Negocios											



Figura 33. Plan de implementación para el proceso propuesto, etapa de proyecciones de negocio.

ETAPA: DEMANDA	ACTIVIDAD	C	ACCIONES	Responsable	Mes n	Mes n+1	Mes n+2	Mes n+3	Mes n+4	Mes n+5	Mes n+6	Mes n+7	Mes n+8	
II.Portafolio	IBP herramienta funcional	B	Actualizar segmentar de productos en IBP (ABC/X,Y,Z)(B) 80% Progress	S&OP										
	IBP herramienta funcional	B	Habilitación del sistema IBP (Planning object de volumen/precio) (B) 70%Progress	S&OP										
	Rediseño de las actividades		M	Confirmar los lanzamientos con casos de negocios adjunto y validar insumos, etc.(M) 30%Progress	Trade									
			M	Comunicar a los equipos para su incorporación (Logística, distribución, plantas)	S&OP									
			M	Optimizar portafolio (Planes de salida y situación de insumos)	S&OP									
			M	Seguimiento de lanzamientos (en base a criterios de éxito/fracaso) (M)	Trade/S&OP									

Figura 34. Plan de implementación para el proceso propuesto, etapa de portafolio de productos.

ETAPA: DEMANDA	ACTIVIDAD	C	ACCIONES	Responsable	Mes n	Mes n+1	Mes n+2	Mes n+3	Mes n+4	Mes n+5	Mes n+6	Mes n+7	Mes n+8
III.Baseline	Aprendizaje y capacitación	A	Diseñar y ajustar modelos estadísticos (fine tuning) 30%Progress	S&OP									
	Aprendizaje y capacitación	A	Entendimiento de modelos y su aplicación según segmentación. (Curva de aprendizaje) (A) 10%Progress	S&OP									
	Aprendizaje y capacitación	A	Potenciar el análisis del entregable para los meses siguientes (2 y 3)	S&OP									
	Aprendizaje y capacitación	A	Generar modelos de pronósticos para precios (A) 25%Progress	S&OP									
	Gestión y Control	M	Revisión detallada de resultados estadísticos generados por modelos (por material; Todos los segmentos) (M)	S&OP									

Figura 35. Plan de implementación para el proceso propuesto, etapa de Baseline.

ETAPA: DEMANDA	ACTIVIDAD	C	ACCIONES	Responsable	Mes n	Mes n+1	Mes n+2	Mes n+3	Mes n+4	Mes n+5	Mes n+6	Mes n+7	Mes n+8
IV.Forecast	Herramienta técnica	A	Realizar Ajuste de precios en IBP	S&OP									
	Gestión y Control	A	Seguimiento y control del MAPE como base para mejorar forecast (A) 5%Progress	S&OP									
	Utilización de la herramienta	A	5.- Separar focos de responsabilidad (en base a segmentación; A,B,C / X,Y)	S&OP-Ventas									
	Capacitación y comunicación	M	Comunicar la importancia de utilizar este resultado como referente para colaborar (resultante de modelos estadísticos avanzados)	S&OP									
	Gestión y Control	M	Revisión exhaustiva de los productos en precio y volumen (Todos los segmentos)	S&OP									

Figura 36. Plan de implementación para el proceso propuesto, etapa de Forecast

ETAPA: DEMANDA	ACTIVIDAD	C	ACCIONES	Responsable	Mes n	Mes n+1	Mes n+2	Mes n+3	Mes n+4	Mes n+5	Mes n+6	Mes n+7	Mes n+8	
V.Input de ventas	Herramienta técnica	A	Generar colaboración importados	S&OP-Ventas										
	Gestión y Control	A	Retroalimentación de la calidad de colaboración.	Ventas										
	Gestión & Estandarización de proceso	B	Colaborar a nivel segmentación (actualmente es a nivel tipo cliente) 40%Progress	S&OP/Ventas										
	Gestión y Control	B	Estandarizar variables base para colaborar(cuantos meses, vta o dda, Fillrate o Mape, IDA). (B) Optimizar esfuerzo destinado a colaborar. Utilizar forecast sugerido s/segmentación (foco A,B,C / Z)	S&OP/Ventas										
	Gestión y Control	B	Estandarizar variables base para colaborar(cuantos meses, vta o dda, Fillrate o Mape, IDA).	S&OP/Ventas										
	Validación y control	M	Medición de volúmenes colaborados según parámetros sugeridos en kickoff	Ventas&Revenue										
		M	Incluir a Revenue como ente validador de precios colaborados. 25%Progress	Ventas										
		M	Fundamentar espíritu de la colaboración (explicación de grandes cambios s/forecast)	Ventas										
Herramienta técnica	M	IBP operativo al 100% (Sin detalles técnicos) 85%Progress	S&OP											

Figura 37. Plan de implementación para el proceso propuesto, etapa de input de ventas.

ETAPA: DEMANDA	ACTIVIDAD	C	ACCIONES	Responsable	Mes n	Mes n+1	Mes n+2	Mes n+3	Mes n+4	Mes n+5	Mes n+6	Mes n+7	Mes n+8
Plan demanda consensuada	Gestión y Control	M	Indicar holguras productivas para intensionar la demanda.	S&OP									
	Gestión y Control	M	Resaltar espíritu Irrestricto de la demanda.	S&OP									
	Gestión & Estandarización	M	Foco en 20/80 (envasados, lanzamientos, etc)	S&OP									
	Integración de Forecast	B	Análisis de Baselines; Forecast v/s colaboración, buscando levantar oportunidades.	S&OP									

Figura 38. Plan de implementación para el proceso propuesto, etapa de plan demanda consensuada.

Supuestos y Restricciones

Los supuestos son factores que para propósitos de la planificación del proyecto se consideran verdaderos, reales o ciertos. A continuación, se presentan algunos los supuestos del proyecto:

Se cuenta con el personal necesario para llevar a cabo las tareas de dirección del proyecto, así como también para dirigir los desarrollos e implementaciones informáticas necesarias para dar cumplimiento a los requerimientos del proyecto.

El proyecto se mantendrá dentro del grupo de proyectos con prioridad en su ejecución para asignar los recursos necesarios.

Las restricciones son factores que limitan la implementación, el rendimiento o las opciones de planificación. Pueden afectar a los objetivos del proyecto o a los recursos que se emplean. A continuación, se presentan las restricciones del proyecto:

El presupuesto no puede exceder el monto que estime la cadena de suministro.

El plazo de implementación no puede exceder el año calendario, de lo contrario habrá que presentar presupuesto para el próximo año.

Riesgos Involucrados

A continuación, se mencionan los riesgos involucrados en la implementación del proyecto:

- Cancelación del proyecto por no estar acorde a nuevos lineamientos

Dimensión: Alineamiento organizacional

Probabilidad de ocurrencia: Bajo

Impacto: Alto

Medida de mitigación: Desarrollar narrativas y ofertas para los actores del proyecto

- Reducción de presupuesto para implementar nuevas tecnologías

Dimensión: Alineamiento organizacional

Probabilidad de ocurrencia: Medio

Impacto: Alto

Medida de mitigación: Implementar con tecnologías alternativas

- Perder el interés y apoyo al proyecto

Dimensión: Involucramiento

Probabilidad de ocurrencia: Bajo

Impacto: Alto

Medida de mitigación: Mantener el proyecto dentro de las prioridades de la empresa

- Postergación o retraso por falta de liderazgo

Dimensión: Liderazgo

Probabilidad de ocurrencia: Medio

Impacto: Alto

Medida de mitigación: Realizar reuniones periódicas y controlar los entregables

- Falta de personal para desarrollo de aplicaciones

Dimensión: Soporte al desempeño

Probabilidad de ocurrencia: Medio

Impacto: Alto

Medida de mitigación: Buscar desarrollador externo

- Pérdida de credibilidad por entrega de información errónea

Dimensión: Soporte al desempeño

Probabilidad de ocurrencia: Medio

Impacto: Alto

Medida de mitigación: Revisión de los datos / Ajuste del modelo

- Cultura organizacional

Dimensión: Involucramiento

Probabilidad de ocurrencia: Alto

Impacto: Alto

Medida de mitigación: Definición de estrategia de gestión del cambio

Principales tareas por desarrollar

Entre las principales tareas a efectuar para la implementación del modelo, se mencionan las siguientes:

- Replicar el modelo desarrollado a nivel SKU, es decir, desagregar el análisis para cada producto y no a la familia de estos como se hizo en el plan piloto
- Generar las consultas y administración de la base de datos necesaria para realizar los pronósticos
- Desarrollar una interfaz de usuario que permita facilitar la ejecución de predicciones
- Capacitación a los planificadores comerciales
- Difundir la nueva herramienta tecnológica

Requerimientos Funcionales y no Funcionales

La implementación del rediseño del proceso incorpora herramientas tecnológicas que deben cumplir ciertos requerimientos. Se debe evaluar una integración a la plataforma SAP IBP para ejecutar directamente en esta tecnología los algoritmos de RNA para el desarrollo y construcción de predicciones.

Se requiere el acceso bajo las normas establecidas por la empresa, con acceso a los sistemas con roles establecidos. El sistema debe ser compatible con las aplicaciones actuales. La interfaz debe ser intuitiva, con lenguaje simple para facilitar la experiencia del planificador o quien esté a cargo del desarrollo.

El área de planificación comercial no utiliza tecnologías ni herramientas de programación en la nube como se propone en los inicios de este plan, se recomienda el uso de tecnologías de la nube para ahorrar costos y potenciar aplicaciones que soporten a este tipo de sistemas. Este formato evita la compra de costosa del hardware que en la mayoría de las veces está sobredimensionado y al poco tiempo queda obsoleta.

Plan de cambio

Todo proyecto tecnológico tiene un impacto en la cultura organizacional de una empresa como también un impacto en las emociones de las personas. Por este motivo, se debe generar una estrategia para lograr una implementación exitosa del proyecto.

Se identifica a los actores del proyecto, su nivel de poder y su relación con el proyecto y la narrativa para poder comunicar adecuadamente y lograr el apoyo necesario. Durante este proceso se debe generar las instancias de comunicación y escuchar a los distintos actores como una manera de fortalecer las conversaciones.

Actor	Nivel de poder	Relación con el proyecto	Narrativa
Gerente Comercial	Alto	Impulsa la dirección hacia la transformación digital con un enfoque en los proyectos que generen una ventaja comparativa con otros actores del mercado	Esta es una solución innovadora dentro de la industria. Su implementación generará una ventaja comparativa con otros actores del mercado.
Gerente Cadena de suministro	Alto	Visualiza una oportunidad de mejora de procesos a través del uso racional de las tecnologías de información que estén alineadas al planteamiento estratégico de la empresa y a proyectos de I+D+i.	Este es un proyecto que está enmarcado dentro de la transformación digital que efectúan las empresas para lograr eficiencia en su operación y una diferenciación con los otros actores del mercado.
Subgerente de Planificación comercial	Medio	Puede cuestionar el proyecto. Sin embargo, deberá entender que la solución es atractiva y estratégica para la empresa y especialmente para su área.	Esta solución apoyará directamente la gestión de la demanda dentro de su área. Será un gran aporte para detectar oportunamente desviaciones que puedan afectar a los objetivos de rentabilidad y ejecución de la empresa.

Tabla 45. Actores directivos involucrados en la decisión de implementar el proyecto y las respectivas narrativas para influenciar en su decisión.

Se debe determinar los hitos del proyecto para programar la difusión del proyecto, las capacitaciones internas que sean necesarias.

Criterios de Aceptación

La aceptación del proyecto y de los entregables estarán sujetos al cumplimiento de los siguientes criterios definidos como estándares dentro de la empresa.

Criterio	Descripción
Compatibilidad	Compatible con tecnologías utilizadas actualmente
Desempeño	Que el margen de error en las predicciones sea menor que los actuales y no exceda +/- 10%
Interoperabilidad	Que permita integración con otros sistemas y que permita exportar e importar datos en formatos estándar (Excel)
Usabilidad	Uso intuitivo, fácil y rápido, con un manual de usuario de ser necesario
Diseño	Diseño consistente con la empresa

Tabla 46. Criterios de aceptación para el proyecto.

12. Evaluación económica del proyecto

Para realizar la evaluación económica del proyecto que se propone, se estimaron primero los costos que éste implica, los cuales se separan en costos de implementación y costos recurrentes o mensuales.

Respecto a los costos de implementación, éstos deben ser pagado una sola vez y se puede ver un resumen de ellos en la siguiente tabla.

Item	Costo
Costo implementación consultora TI	\$10.000.000
Licencias de cuentas Datacamp	\$1.530.000
TOTAL	\$11.530.000

Tabla 47. Costos de implementación para llevar a cabo el proyecto.

El primer ítem que aparece es el costo que cobraría una consultora especializada en la implementación y desarrollo de proyectos de TI, la idea es a través de esta consultoría escalar la solución de manera que se pueda pronosticar todos los SKU de Agrosuper mediante una sola ejecución del algoritmo, y no tener que hacerlo de manera manual uno por uno ya que representaría muchos inconvenientes y costos operativos para la subgerencia de planificación comercial. Se cotizaron en [42-43] los precios de un proyecto con características similares al planteado Este es un proceso que demorará aproximadamente tres meses según las consultas realizadas, en el cual se considera la implementación, las configuraciones e integración de todas las funcionalidades esperadas, el desarrollo de las especificaciones que se deban hacer al software y una capacitación para el uso de este para los trabajadores del área.

También se incluye la contratación de cuentas Datacamp para el equipo con el objetivo de capacitar en los entornos y lenguajes de programación de Python, pues si se quiere escalar a implementar estos modelos de redes neuronales artificiales, los colaboradores del equipo deben manejarse con temas de programación y algoritmos en este u otro lenguaje.

Considerando todo lo anterior, el costo total de implementación del rediseño es de \$11.530.000 aproximadamente.

Respecto a los costos recurrentes, éstos deben ser pagados cada mes y se puede ver un resumen de ellos en la tabla siguiente.

Item	Costo Mensual
Remuneración Ingeniero de desarrollo TI	\$1.500.000
Costos de apoyo al proceso de demanda	\$1.000.000
Soporte de TI	\$250.000
TOTAL	\$2.750.000

Tabla 48. Costos mensuales para llevar a cabo el proyecto.

El primer ítem que se refiere a la remuneración de un consultor de apoyo, el cual debe ser ingeniero con un manejo en desarrollo de algoritmos de redes neuronales artificiales y una base sólida con el trabajo con bases de datos que será contratado por la empresa para apoyar la comunicación con la consultora y resolver todos los problemas que puedan surgir durante la implementación de la solución, incluyendo los temas más propios de la empresa, el sueldo tendrá un ajuste anual que será un aumento de un 5% de crecimiento respecto al año anterior, las funciones y tareas serán algunas de las siguientes:

- Diseñar y desarrollar arquitectura de base de datos, estructuras de datos, tablas, diccionarios y convenciones de nomenclatura para el proyecto
- Diseñar, construir, modificar, integrar, implementar y probar los sistemas de gestión de bases de datos
- Realizar investigaciones y prestar asesoramiento sobre la selección, aplicación y ejecución de herramientas de gestión de bases de datos
- Desarrollar y aplicar políticas, documentación, normas y modelos de administración de datos
- Desarrollar políticas y procedimientos para el acceso, uso, copia de seguridad y recuperación de las bases de datos
- Establecer el funcionamiento y mantenimiento preventivo de las copias de seguridad, procedimientos de recuperación, y reforzar la seguridad y la integridad de los controles
- Escalar y desarrollar el algoritmo de predicción de RNAs de manera que se puedan realizar muchas predicciones de SKUs en la misma ejecución del programa, o bien, de manera simultánea, aumentando la productividad y efectividad.

El segundo ítem corresponde a un presupuesto que cada mes se gastará con el fin de invertirlo en mejorar el proceso de demanda, ya sea con horas de trabajo extra operativo de los colaboradores del área, capacitaciones extras, costos de administración, entre otros que se puedan tener durante la implementación. Cabe destacar que este presupuesto aumentará en un 10% cada año.

Finalmente, se incluye un costo de soporte, el cual es entregado por la consultora especializada que implementa el algoritmo. Este costo es de \$250.000 durante los primeros años, ya que incluye el soporte al software que se quiere programar, y a la resolución de errores en las modificaciones que se hagan. De esta forma, el costo mensual total del rediseño, durante el primer año corresponde a \$2.750.000.

Respecto a los beneficios, éstos pueden ser de corto, mediano y largo plazo. En el corto plazo, los beneficios tienen relación con aumentar las ventas a través de una menor cantidad de ventas perdidas, y, además, disminuir los costos que tiene para la empresa los errores que se cometen en la gestión de planificación y en su corrección.

En el mediano y largo plazo, se pretende conseguir una mejor ejecución en la planificación que se establece mensualmente, intentar bajar considerablemente los errores de pronóstico, contar con mayor disponibilidad, con lo cual se espera una mejora en la planificación de la cadena de suministro que se traduzca también en una mejora en la ejecución del plan con un aumento de la cantidad de ventas de la empresa.

Dada la dificultad para traducir los beneficios antes descritos a ganancias económicas, se evalúa la conveniencia del rediseño mediante los resultados del plan piloto mostrado previamente. Se

considera lo siguiente para el cálculo de los beneficios económicos, una mejora de un 1% en los pronósticos de demanda, mejorará en un 0,3% en el fill rate, o bien, el nivel de servicio. Esto debido a que, al pronosticar y planificar mejor la demanda y los planes comerciales, conlleva a tener una mejor ejecución e incluso una mejor distribución de productos, aumentando la disponibilidad de estos y evitando quiebres de stock para ser correctamente comercializados. Vale decir que, este escenario del 0,3% es un caso base, en donde también se analizarán escenarios pesimista y optimista, en donde el nivel de servicio mejoraría en un 0,25% y 0,4% respectivamente.

Por tanto, si un producto, por ejemplo, fue pronosticado con un mejor desempeño con el modelo de RNA en un 3% respecto a los modelos de SAP IBP, entonces se considera un nivel de servicio aumentado en un 0,9% y por tanto se calculan esos kilogramos de demanda por su respectivo precio y se considera como una ganancia económica del proyecto. Con esto en mente, se hizo el cálculo de la cantidad de kilos de cada material mejor pronosticados y se llegó a la conclusión que 893.988 kilos de demanda en los distintos sectores (cerdo, aves y procesados) fueron mejor provistos por el modelo de RNA, por tanto, con este valor se calculó cuantos kilogramos más se podrían haber vendido aumentando el nivel de servicio mencionado y se llega a la estimación de \$4.861.946 en ganancias económicas en el escenario base para las 5 semanas de prueba, por lo que este valor se escala proporcionalmente a un monto anual considerando que un año cuenta con 52 semanas. Además, se considera para el calculo del flujo de caja un aumento en las cantidades mejor pronosticadas de un 0,5% cada año a medida que se van instaurando las mejoras propuestas en el proyecto. Y se evaluará el proyecto en un horizonte de 5 años.

Por otro lado, para poder analizar los indicadores de rentabilidad, fue necesario calcular la tasa de descuento del sistema, por lo que se siguió la siguiente fórmula del modelo CAPM [45-47] para su cálculo:

$$R_d = r_{LR} + \beta(r_m - r_{LR}) + RP$$

Donde las diferentes siglas representan:

r_{LR} : Tasa libre de riesgo

β : Cantidad de riesgo

r_m : Retorno esperado de portafolio

RP : Riesgo del país

La siguiente tabla muestra los valores de estos indicadores y la tasa de descuento final:

r_{LR}	2,23%
B	1,06
r_m	13,19%
RP	3,14%
Tasa de descuento	16,94%

Tabla 49. Cálculo de la tasa de descuento según el modelo CAPM.

A continuación, se presenta el flujo de caja realizado:

	Año 0	Año 1	Año 2	Año 3	Año 4	Año 5
	Montos en miles de pesos					
Aumento de ventas		50.554	50.807	53.348	56.015	58.816
Ingresos		50.554	50.807	53.348	56.015	58.816
Sueldos		18.000	18.900	19.845	20.837	21.879
Trabajo operativo		12.000	12.120	12.241	12.364	12.487
Soporte		3.000	3.000	3.000	3.000	3.000
Costos variables		-33.000	-34.020	-35.086	-36.201	-37.366
Licencias de Datacamp		1.530	1.576	1.623	1.672	1.722
Costos fijos		-1.530	-1.576	-1.623	-1.672	-1.722
Utilidad antes de impuesto		16.024	15.211	16.638	18.142	19.727
Impuesto (27%)		4.327	4.107	4.492	4.898	5.326
Utilidad después de impuesto		11.698	11.104	12.146	13.244	14.401
Flujo operacional		11.698	11.104	12.146	13.244	14.401
Inversión	-10.000					
Flujo de capitales	-10.000					
Flujo de caja	-10.000	11.698	11.104	12.146	13.244	14.401

Tabla 50. Flujo de caja para los primeros 5 años de implementación del proyecto.

Luego, con este flujo se calcula el VAN en miles de pesos chilenos y la TIR para obtener los indicadores de rentabilidad del proyecto, se analizan los resultados por escenarios (pesimista, base y optimista) a través de la variación en el nivel de servicio descrito previamente, dando como resultado los siguientes valores:

	Pesimista	Base	Optimista
VAN	\$-6.111	\$25.129	\$60.641
TIR	-17%	115%	240%

Tabla 51. Indicadores de rentabilidad calculados para los escenarios definidos en base al flujo de caja.

Por tanto, en base a estos indicadores es posible comentar que el proyecto representa una oportunidad rentable, ya que el valor de su VAN es de 25 millones de pesos aproximadamente mientras que su TIR es de un 115%. Además, es claro observar que el proyecto es muy sensible a cambios en el porcentaje de nivel de servicio que se espera lograr con la disminución de los errores de pronóstico. En definitiva, los costos que implica el proyecto son menores al compararlos con los beneficios potenciales que traerá para la empresa la implementación de esta metodología, por lo cual es económicamente conveniente implementar el rediseño

13. Conclusiones y discusión

El modelo de S&OP en la actualidad es de vital importancia para Agrosuper, donde el mercado es tan volátil y la demanda cambia constantemente de acuerdo con la necesidad de los consumidores, ya que este es el encargado de alinear a las diferentes áreas para lograr el resultado esperado por la estrategia de la Dirección General mediante planes tácticos y operativos que debe ejecutar la cadena de suministro en la operación diaria.

En el proceso y en la planificación de la demanda la coordinación y el liderazgo son claves para garantizar una correcta ejecución de todo el proceso, la información y el flujo de esta debe ser eficiente para lograr los consensos en las diferentes áreas involucradas. El S&OP es una guía de ruta para la ejecución de planes tácticos (mensual) y operativos (semanal), pero no es el entregable final a las diferentes áreas el que garantiza el resultado, sino el mismo consenso y seguimiento al que se llega y se lleva durante su construcción y ejecución. El S&OP se hace imprescindible por el constante cambio en la demanda, por lo que es sumamente importante fortalecer y agregarle valor a este. Se puede concluir que el S&OP se convierte en una herramienta de gestión donde la participación y compromiso de todas las áreas es la que finalmente le dará éxito a su implementación.

En resumen, el proceso de S&OP debe cumplir las siguientes características:

- S&OP constituye una serie de pasos establecidos con frecuencia mensual/semanal y no una reunión de gerencia.
- S&OP es un proceso de alto nivel, donde los directivos toman las decisiones de rentabilidad de la compañía.
- S&OP trabaja a nivel agregado, con familias de producto y subfamilias.
- S&OP es un proceso con visión a corto, a mediano e incluso a largo plazo.
- El foco de S&OP es el futuro, sin embargo, sus indicadores son revisados periódicamente a fin de que los problemas sean visibles, y se estimule el mejoramiento continuo.
- S&OP es un proceso transversal en toda la compañía, la dirección del equipo toma decisiones del interés de toda la compañía, que no necesariamente coinciden con los intereses individuales de las áreas o departamentos.
- El plan aprobado por el grupo de S&OP, es el plan que debe ejecutarse, y cualquier cambio en un área amerita el cambio del plan y su nueva aprobación y divulgación.
- Los planes de S&OP deben orientar el futuro del negocio, y de ser necesario es fundamental la simulación de escenarios ("what if") que permitan desarrollar planes de contingencia basados en los diferentes casos.

Se recomienda al área de planificación comercial de la empresa revisar la posibilidad de usar el algoritmo de redes neuronales artificiales, ya que es una herramienta que permite a planeadores y líderes del proceso tener una visual más integral y clara de la demanda. El impacto y relevancia que potencialmente tiene este modelo para apoyar en las decisiones que se toman se debe a que entrega otra perspectiva dando una visión más holística de este complejo proceso de estimación.

Además, debido a la variación aleatoria que siempre existe, se debe considerar que, al realizar pronósticos de demanda, nunca se podrán hacer predicciones sin errores, lo que no significa que no se deban realizar estimaciones de demanda, ya que de todas formas son un aporte para la planificación de la empresa. El nuevo modelo es una herramienta adicional muy interesante, que

dejará grandes aprendizajes y conocimiento para todos aquellos que trabajan en el ámbito de la demanda y la predicción.

Finalmente podemos concluir que, aunque el proceso de S&OP y el de pronóstico de demanda en Agrosuper es organizado, tiene responsables y un cronograma detallado, falla en algunos temas relevantes respecto a la concepción de este, que hacen que no sea una herramienta de gestión donde es el consenso el que define todos los entregables, sino que cada área revisa sus cifras y al final se entrega un resultado para tratar de llegar a un consenso. No debe ser un documento al final de mes el que defina los planes a seguir, sino que el mismo proceso y el consenso en cada etapa es el que debiera definir estos planes, y el entregable al final del proceso simplemente condensa estos acuerdos y garantiza que se cumplan durante el mes; sólo de esta manera se podrá lograr la alineación esperada y mejorar el subproceso de demanda, donde la estrategia irá de la mano de los planes operativos, que garantizarán el resultado deseado.

Con base en información histórica, en el presente trabajo de título, se realizó un análisis exhaustivo de la previsión de demanda de productos de consumo masivo. Se analizaron las series de tiempo caracterizándolas en función de las toneladas demandadas. La previsión de la demanda fue estimada usando los métodos de previsión clásicos provistos por SAP IBP: Auto-ARIMA, Auto-SARIMA, Alisamiento exponencial automático, Regresión lineal temporal, y usando métodos de aprendizaje automático con base en Redes Neuronales Artificiales mediante la programación realizada sólo para este trabajo de título en entornos de Google Colab.

El estudio, plan y comparación de efectividad que se hizo, mostró que los pronósticos de la demanda del 31% de los casos fueron mejor previstos por el algoritmo de redes neuronales artificiales, estos errores en la predicción fueron medidos por el indicador de MAPE y por tanto en esos casos este fue menor que aquellos algoritmos provistos por SAP IBP. Además, se evidenció que los sectores de cerdo y pollo el algoritmo de RNA obtuvo un mejor desempeño relativo a los otros métodos de previsión, por tanto, en estos negocios se divisa la oportunidad de utilizarlo a futuro.

A pesar de tener resultados satisfactorios en algunos casos, no se puede demostrar que la previsión con algoritmos de aprendizaje automático es mejor que la previsión usando métodos más tradicionales, además que la especificación de las redes y la fase de entrenamiento constituyen procesos muy sensibles y de cierta manera relativamente críticos. Saber que cierto método sofisticado no es tan preciso como uno mucho más simple es perturbador desde un punto de vista científico, ya que el primero requiere una gran cantidad de experiencia académica y mucho tiempo en la computadora para ser aplicado.

En las comparaciones de los métodos estadísticos y de RNA informados en este documento, debe quedar claro que los resultados pueden estar relacionados con el conjunto de datos específico que se han utilizado.

Respecto a las hipótesis y objetivos planteadas en el inicio de este trabajo, se pensaba que los métodos de predicción de aprendizaje automático tenían una precisión superior simplemente por su sofisticación y su elegancia matemática. Ahora, es obvio que su valor debe ser probado empíricamente de manera objetiva e indiscutible a través de investigaciones rigurosas y a gran escala. Por lo tanto, cuando se trata de artículos que proponen nuevos métodos de aprendizaje automático o formas efectivas de usarlos, se deben exigir comparaciones con métodos alternativos

o al menos puntos de referencia y exigir que los datos de los artículos que se publican estén disponibles para quienes deseen replicar los resultados para validarlas.

Además de las pruebas empíricas, se necesita trabajo de investigación y capacitaciones para ayudar a los planificadores comerciales a comprender cómo se generan los pronósticos de los métodos de RNA (este es el mismo problema con todos los modelos de IA cuya salida no se puede explicar). Obtener números de una caja negra no es aceptable para los profesionales que necesitan saber cómo surgen los pronósticos y cómo pueden ser influenciados o ajustados para llegar a predicciones viables.

Una última preocupación igualmente importante es que, además de los pronósticos puntuales, los métodos de RNA también deben ser capaces de especificar la incertidumbre en torno a ellos o, alternativamente, proporcionar intervalos de confianza. En la actualidad, el tema de la incertidumbre no se ha incluido en la agenda de investigación, dejando un enorme vacío que debe llenarse ya que estimar la incertidumbre en las predicciones futuras es tan importante como los propios pronósticos.

En este punto, se pueden hacer las siguientes sugerencias / especulaciones, que deben ser verificadas empíricamente, sobre el camino a seguir con respecto a los métodos de aprendizaje automático, mientras que estas pueden enriquecerse con temas de investigación futuros propuestos:

- Obtener más información sobre los valores futuros desconocidos de los datos en lugar de sus valores pasados y enfocar la optimización / aprendizaje en dichos valores futuros tanto como sea posible.
- Desestacionalizar los datos antes de utilizar métodos de RNA. Esto resultará en uno más simple, reduciendo el tiempo computacional requerido para llegar a pesos óptimos y, por lo tanto, aprender más rápido.
- Evitar el ajuste excesivo, ya que no está claro si los modelos de RNA pueden distinguir correctamente el ruido del patrón de los datos.
- Automatizar el preprocesamiento y evitar las decisiones adicionales requeridas por parte de los planificadores.
- Permitir la estimación de la incertidumbre para los pronósticos puntuales y proporcionar información para la construcción de intervalos de confianza alrededor de dichos pronósticos.

Aunque la conclusión de este trabajo de que la precisión de pronóstico de los modelos de RNA es menor que la de los métodos estadísticos como el de auto-SARIMA puede parecer decepcionante, se piensa extremadamente positivo sobre el gran potencial de los modelos de RNA para las aplicaciones de pronóstico. Claramente, se necesita más trabajo para mejorar dichos métodos, pero lo mismo ha ocurrido con todas las nuevas técnicas, incluidos los complejos métodos de pronóstico que han mejorado su precisión considerablemente con el tiempo.

Además, al evaluar económicamente este trabajo se llega a que representa una oportunidad rentable incorporar estos algoritmos al apoyo de las decisiones en la cadena de suministro de Agrosuper. En la medida que se desarrolla un proyecto según una especificación inicial, se van vislumbrando los posibles trabajos futuros o mejoras, pues no será posible incorporar todo en la actual versión por restricciones de alcance, costo y tiempo.

Son varios los trabajos que deberán evaluarse como nuevos proyectos pues se requerirá una cantidad de recursos para llevarlos a cabo este proyecto, sin embargo, este riesgo de inversión puede significar en una ventaja competitiva a futuro para este mercado. Es importante comentar que hace quince años no se creía que existiría asistencia personal en nuestros teléfonos móviles para comprender y hablar en idiomas naturales, traducciones automáticas en la web, casos como algoritmos de reconocimiento de expresiones faciales, por mencionar algunos ejemplos. No hay ninguna razón por la que no se pueda lograr el mismo tipo de avances con los métodos de aprendizaje automático (RNAs) aplicados a la previsión.

Sin embargo, debemos darnos cuenta de que aplicar la inteligencia artificial a la predicción es bastante diferente a hacerlo en los juegos o en el reconocimiento de imágenes y voz y puede requerir diferentes algoritmos especializados para tener éxito. A diferencia de otras aplicaciones, el futuro nunca es idéntico al pasado y el entrenamiento de los métodos de IA no puede depender exclusivamente de él.

14. Bibliografía

- [1] OCDE/FAO (2020). *Cómo alimentar al mundo en 2050*, FAO. [en línea] Consultado el 18 de Octubre de 2021. Sitio web: (www.fao.org)
- [2] OCDE/FAO (2017), “Carne”, en *OCDE-FAO Perspectivas Agrícolas 2017-2026*, OECD Publishing, París. Consultado el 18 de Octubre de 2021. [en línea] Sitio web: (<https://data.oecd.org/agroutput/meat-consumption.htm>)
- [4] OCDE. (2018). *Los desafíos de Chile en inocuidad y calidad agroalimentaria al 2030: Una propuesta institucional de gestión*, 2018.
- [5] Burgman MA. (2015). *Juicios confiables: cómo sacar lo mejor de los expertos*. Prensa de la Universidad de Cambridge; 2015.
- [6] DJ Crawford-Brown. *Decisiones ambientales basadas en riesgos: métodos y cultura*. Springer, Boston, MA; 1999.
- [7] Cooke RM. *Expertos en incertidumbre: opinión y probabilidad subjetiva en ciencia*. Serie de Política Científica y Ética Ambiental. Prensa de la Universidad de Oxford; 1991.
- [8] Mandel DR, Barnes A. *Precisión de los pronósticos en inteligencia estratégica*. *Procedimientos de la Academia Nacional de Ciencias*. 2014; 111 (30): 10984–10989.
- [9] Zhang, Peter: “*Neural Networks in Business Forecasting*”, 1ª edición, Idea Group Publishing, E.U.A. 2004.
- [10] SAP Integrated Business Planning (2021) SAP IBP. Recuperado el 18 de Octubre del 2021. Sitio web: (https://help.sap.com/viewer/product/SAP_INTEGRATED_BUSINESS_PLANNING/2108/en-US?task=whats_new_task)
- [11] Zhang Peter, Patuwo Eddy and Hu Michael. (1998). *Forecasting With Artificial Neural Networks: The State of the Art*. *International Journal of Forecasting*, 14, 35-62.
- [12] Makridakis Spyros, Spiliotis Evangelos and Assimakopoulos Vassilis. (2018). *Statistical and Machine Learning forecasting methods: Concerns and ways forward*. *PLoS ONE*, 13.
- [13] Makridakis Spyros, Spiliotis Evangelos and Assimakopoulos Vassilis. (2018). *The M4 Competition: Results, findings, conclusion and way forward*. *International Journal of Forecasting*, 34, 802-808. 62

- [14] Fry Chris, Brundage Michael. (2019). The M4 forecasting competition - A practitioner's view. *International Journal of Forecasting*, 36, 5.
- [15] Redd A., Khin K. and Marini A. (2019). Fast ES-RNN: A GPU Implementation of the ES-RNN Algorithm. *ArXiv*, abs/1907.03329.
- [16] M. Casey Brace, J. Schmidt and M. Hadlin. (1991). Comparison of the forecasting accuracy of neural networks with other established techniques. *Proceedings of the First International Forum on Applications of Neural Networks to Power Systems*, Seattle, WA, USA, 31-35.
- [17] Chakraborty Kanad, Mehrotra Kishan, Mohan Chilukuri and Ranka Sanjay. (1992). Forecasting the Behavior of Multivariate Time Series Using Neural Networks. *Neural Networks*, 5, 961-970.
- [18] Foster W.R., Collopy Fred and Ungar Lyle. (1992). Neural network forecasting of short, noisy time series. *Computers and Chemical Engineering*, 16, 293-297.
- [19] Hann Tae and Steurer Elmar. (1996). Much ado about nothing? Exchange rate forecasting: Neural networks vs. linear models using monthly and weekly data. *Neurocomputing*, 10, 323-339.
- [20] Fulcher Ben. (2018). *Feature-Based Time-Series Analysis*. CRC Press.
- [21] Wang Xiaoyue, Ding Hui, Trajcevski Goce, Scheuermann Peter and Keogh Eamonn. (2013). Experimental Comparison of Representation Methods and Distance Measures for Time Series Data. *Data Mining and Knowledge Discovery*, 26, 275309.
- [22] Hochreiter Sepp and Schmidhuber Jürgen. (1997). Long Short-term Memory. *Neural computation*, 9, 1735-1780.
- [23] Aileen Nielsen. (2019). *Practical Time Series Analysis: Prediction with Statistics and Machine Learning*. O'Reilly.
- [24] Denis Kwiatkowski, Peter C.B. Phillips, Peter Schmidt and Yongcheol Shin. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?. *Journal of Econometrics*, 54 (1-3), 159-178.
- [25] Brockwell P. J. and Davis R. A. (2016). *Introduction to time series and forecasting (3rd ed)*. New York, USA, Springer.
- [26] Wei YH, Wang BS, Hao SS. *Predicción estadística y toma de decisiones*. Chengdu: Prensa de la Universidad Southwest Jiaotong, 2014

[27] Ren S, Cui HB. Aplicación del análisis de series de tiempo SARIMA en el pronóstico de impuestos: tome la provincia de Guizhou como ejemplo. Revista de la Universidad de Hubei (Ciencias Naturales), 2021, 43 (1): 80–85.

[28] Multimodelo de selección e inferencia de modelos de Burnham KP y Anderson DR: un enfoque teórico práctico. 2ª edición, Springer-Verlag, Berlín. 2002

[29] Bengio Y., Simard Patrice and Frasconi Paolo. (1994). Learning longterm dependencies with gradient descent is di-cult. IEEE transactions on neural networks, 5, 157-166.

[30] Hochreiter Sepp and Schmidhuber Jürgen. (1997). Long Short-term Memory. Neural computation, 9, 1735-1780.

[31] KDNuggets.com, (2014, Agosto), “What programming/statistics languages you used for an analytics / data mining / data science work in 2014?” [En línea], Disponible en: <http://www.kdnuggets.com/polls/2014/languages-analytics-data-mining-data-science.html>

[32] J. Gallardo, “Metodología para la Definición de Requisitos en Proyectos de Data Mining (ER-DM)”, Tesis Doctoral, Universidad Politécnica de Madrid, Madrid, ES, 2009.

[33] C. Shearer, “The CRISP-DM model: The new blueprint for data mining,” Journal of Data Warehousing, Vol. 5, No. 4, pp 13-22, Otoño, 2000.

[34] P. Chapman, J. Clinton, R. Kerber, T. Khabaza, T. Reinartz, C. Shearer y R. Wirth, “CRISP-DM 1.0: Step-by-step data mining guide”, USA: 2000.

[35] Oscar Barros (2009). Ingeniería de negocios Diseño integrado de negocios, procesos y aplicaciones TI. Santiago de Chile.

[36] Oscar Barros (2016). Ingeniería de Negocios: Diseño Integrado de Servicios, sus Procesos y Apoyo TI. Versión Kindle.

[37] Richard B. Chase, F. Robert Jacobs, Nicholas J. Aquilano (2009). Administración de operaciones. Producción y cadena de suministros. México D.F. 12º Edición, Editorial McGraw-Hill.

[38] Stephen A. White, Derek Miers (2009). Guia de Referencia y Modelado BPMN: Comprendiendo y utilizando BPMN. Edición digital, por Future Estrategies Inc., Book Division.

[39] Oscar Barros (2000). Rediseño de Procesos de Negocios Mediante el Uso de Patrones. Santiago de Chile. Dolmen Ediciones.

[40] Agrosuper. (2021). Reporte Integrado de Agrosuper 2020. [en línea] Recuperado el 9 de septiembre de 2021. Sitio web: <<https://www.agrosuper.cl/wp-content/uploads/2021/04/Reporte-Integrado-Matriz-Agrosuper-2020-web.pdf>>

[41] Agrosuper. (2020) Reporte Integrado de Agrosuper 2019. [en línea] Recuperado el 9 de septiembre de 2021. Sitio web: <<https://www.agrosuper.cl/wp-content/uploads/2020/07/Reporte-Integrado-2019-web-2.pdf>>

[42] Software ERP Defontana Chile (2021). Planes y precios | Defontana Chile. [en línea] Recuperado el 9 de diciembre de 2021, de Defontana. Sitio web: <https://www.defontana.com/cl/productos/erp/>

[43] Software ERP Defontana Chile (2021). Integraciones | Defontana Chile. [en línea] Recuperado el 9 de diciembre de 2021, de Defontana. Sitio web: <https://www.defontana.com/cl/productos/integraciones/>

[44] Tusalarario.org (2021). Tusalarario.org/Chile - Comparador Salarial, Conoce el salario neto y bruto de tus colegas en Chile [online]. Recuperado el 9 de diciembre de 2021, de Tusalarario.org. Sitio web: <https://tusalarario.org/chile/main/salario/Comparatusalarario?job-id=3343010000000#/>

[45] Indicadores Larrain Vial. (2021). "Indicadores de mercado" [en línea] Recuperado el 18 de noviembre de 2021. Sitio web: <https://larrainvial.finmarketslive.cl/www/v2/index.html?mercado=chile>

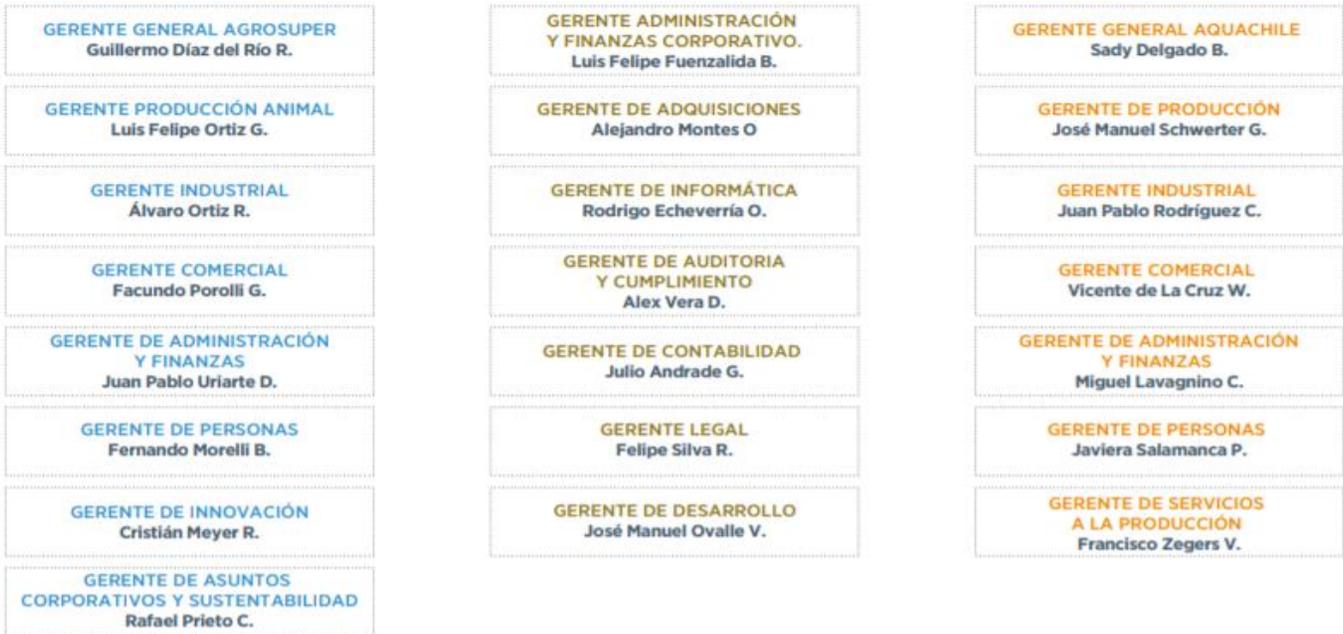
[46] Betas. (2021) "Betas por sector (US)" [en línea] Recuperado el 18 de noviembre de 2021. Sitio web: http://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/Betas.html

[47] S&P CLX IGPA datos históricos. (2021) "SP CLX IGPA" [en línea] Recuperado el 18 de noviembre de 2021. Sitio web: <https://es.investing.com/indices/igpa-historical-data>

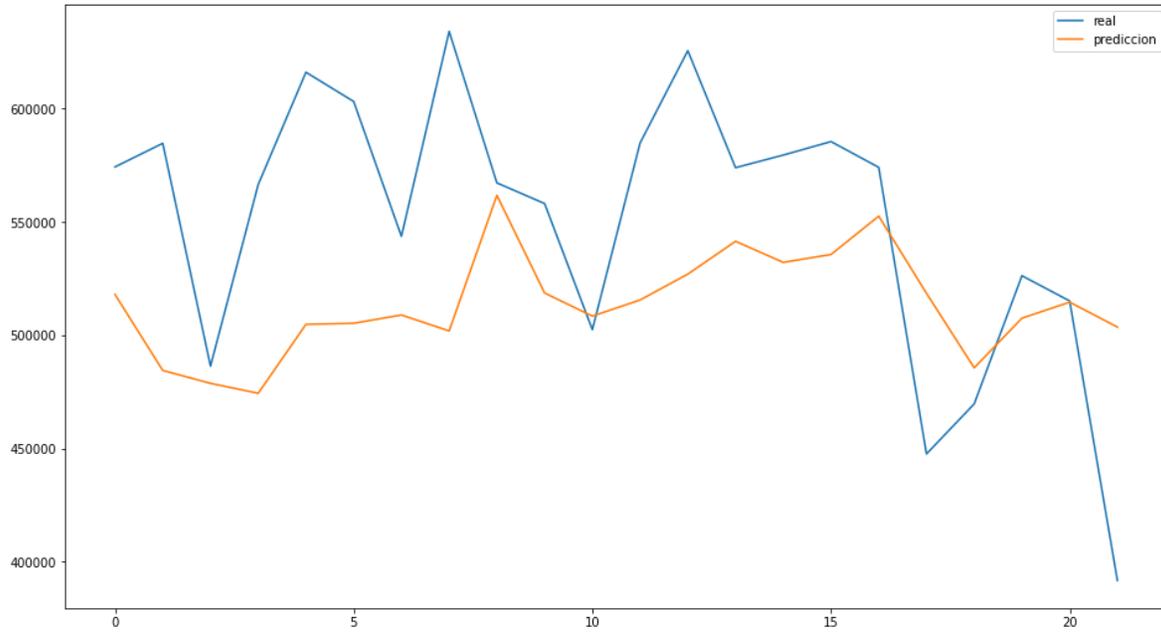
[48] Zeng N, Zhang H, Song B, Liu W, Li Y, Dobaie AM. Facial expression recognition via learning deep sparse autoencoders. Neurocomputing. 2018;273(Supplement C):643–649

MATRIZ
AGROSUPER

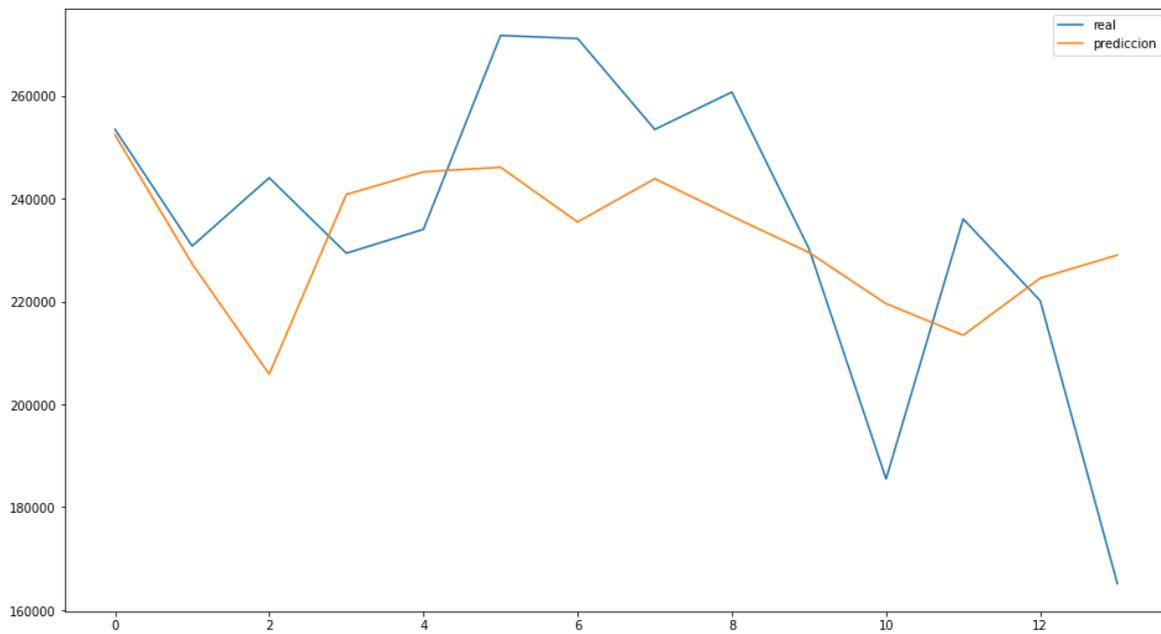
GERENTE GENERAL
CORPORATIVO
José Guzmán V.



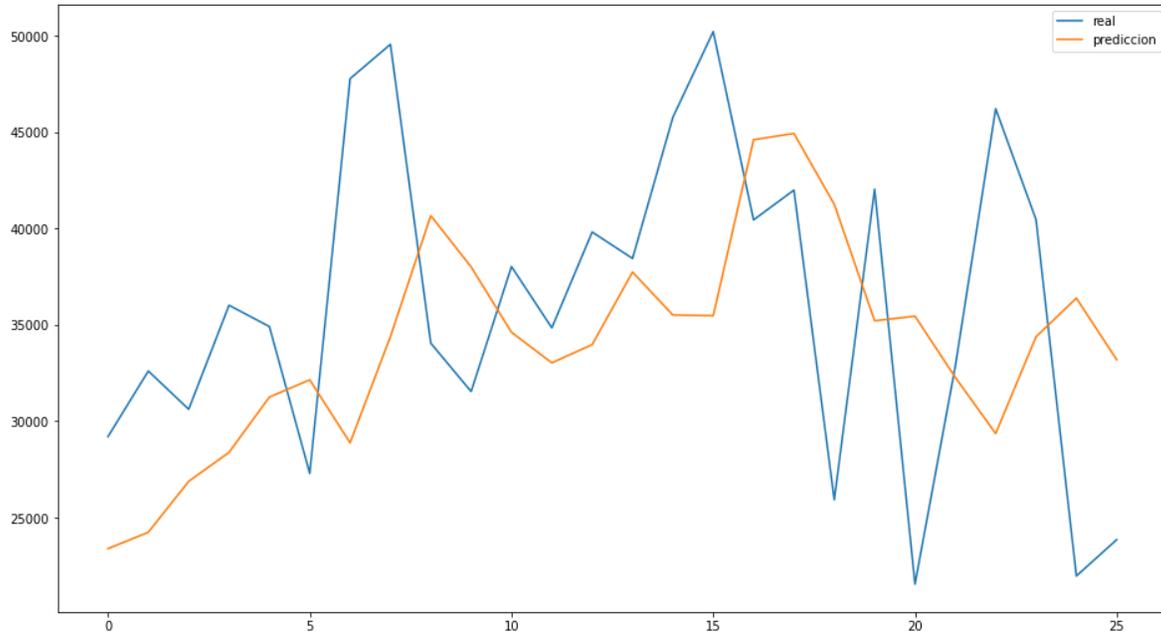
Anexo b. Organigrama de la Matriz de Agrosuper. Fuente: Agrosuper.



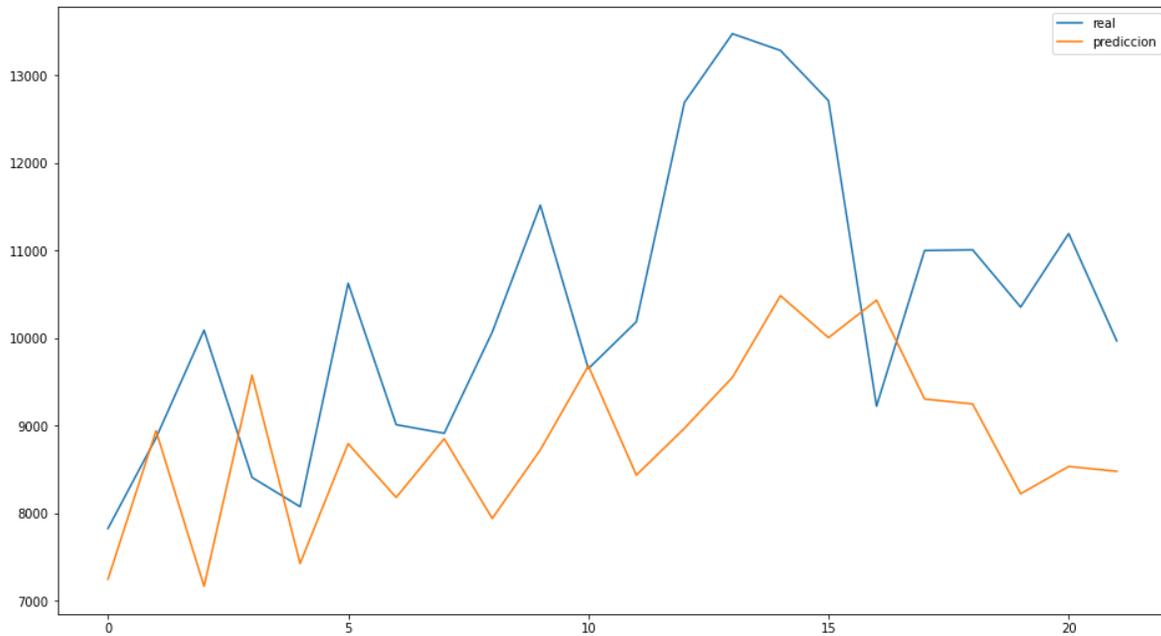
Anexo c. Predicción vs datos reales de la categoría de chuleta de cerdo con un mape de 10,02% de error.



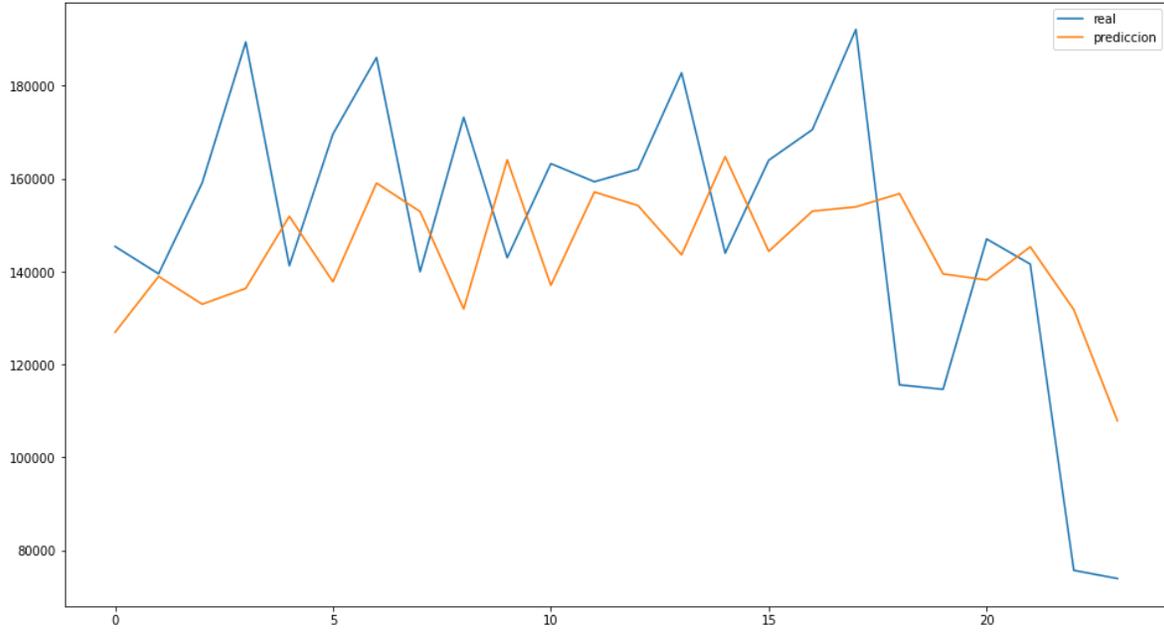
Anexo d. Predicción vs datos reales de la categoría de costillas-pecho de cerdo con un mape de 9,42% de error.



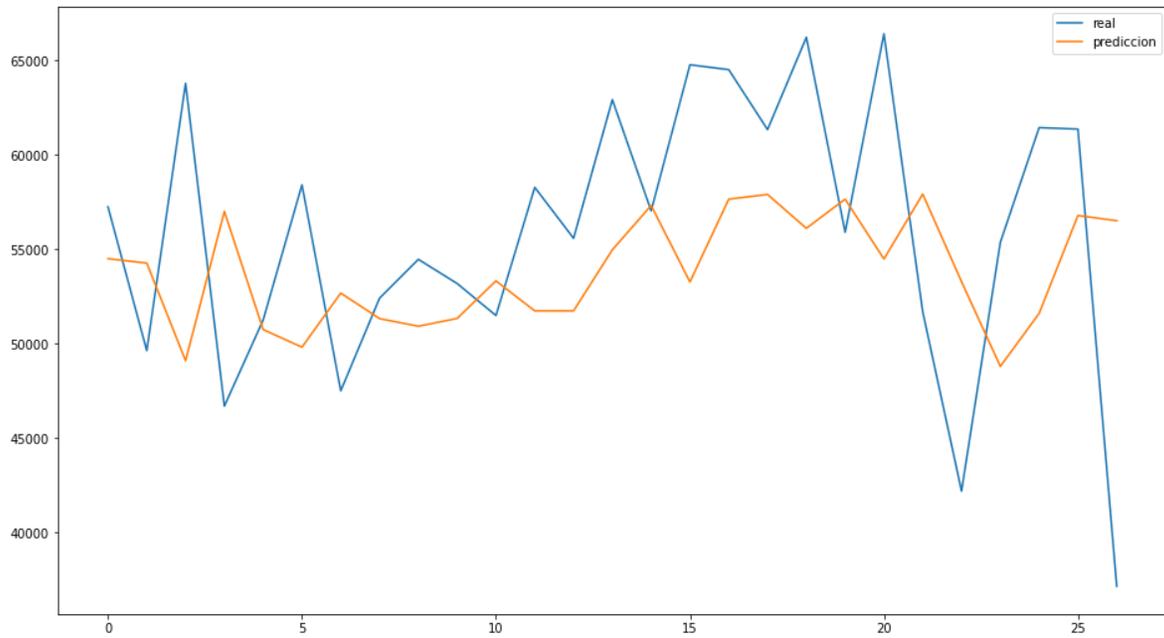
Anexo e. Predicción vs datos reales de la categoría de cueros de cerdo con un mape de 23,63% de error.



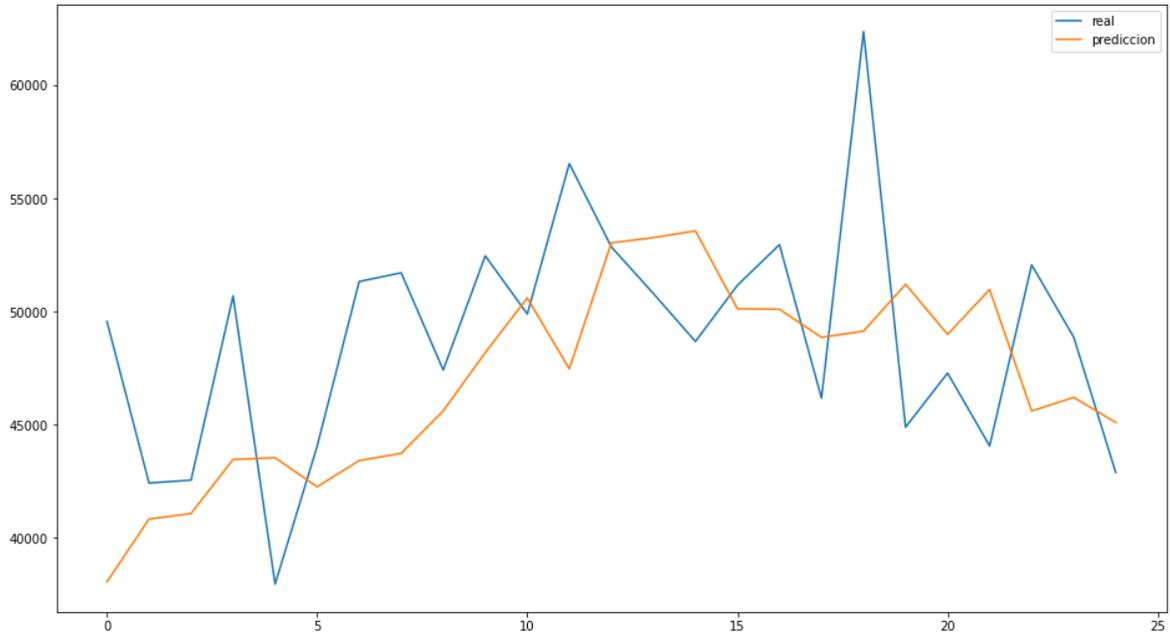
Anexo f. Predicción vs datos reales de la categoría de filete de cerdo con un mape de 16,08% de error.



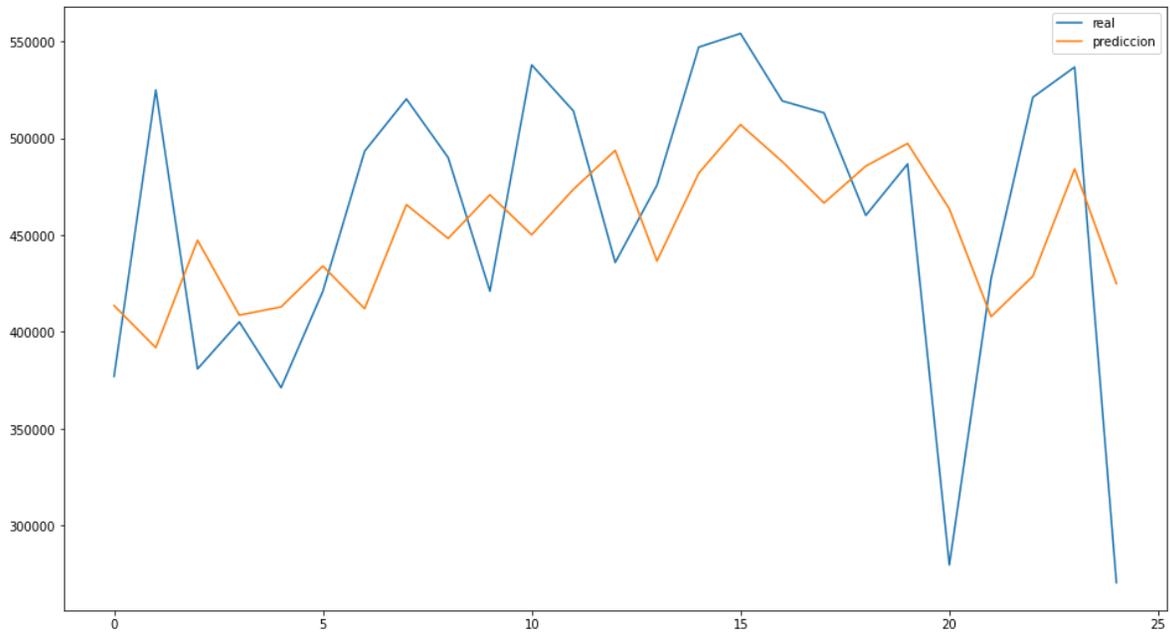
Anexo g. Predicción vs datos reales de la categoría de grasa de cerdo con un mape de 18,02% de error.



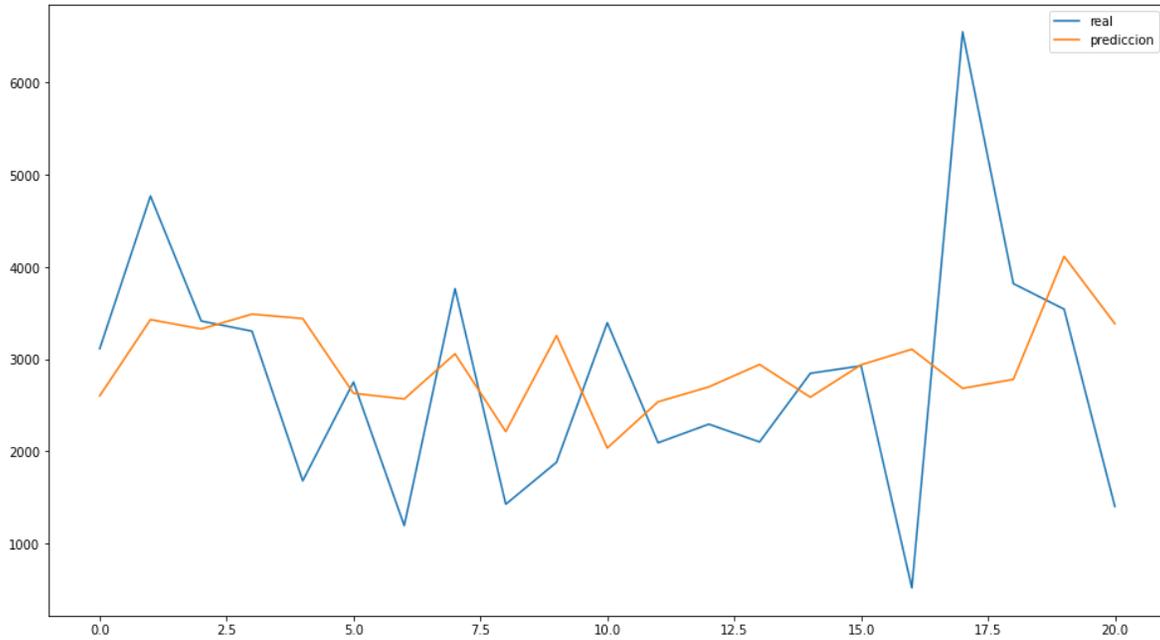
Anexo h. Predicción vs datos reales de la categoría de huesos de cerdo con un mape de 12,19% de error.



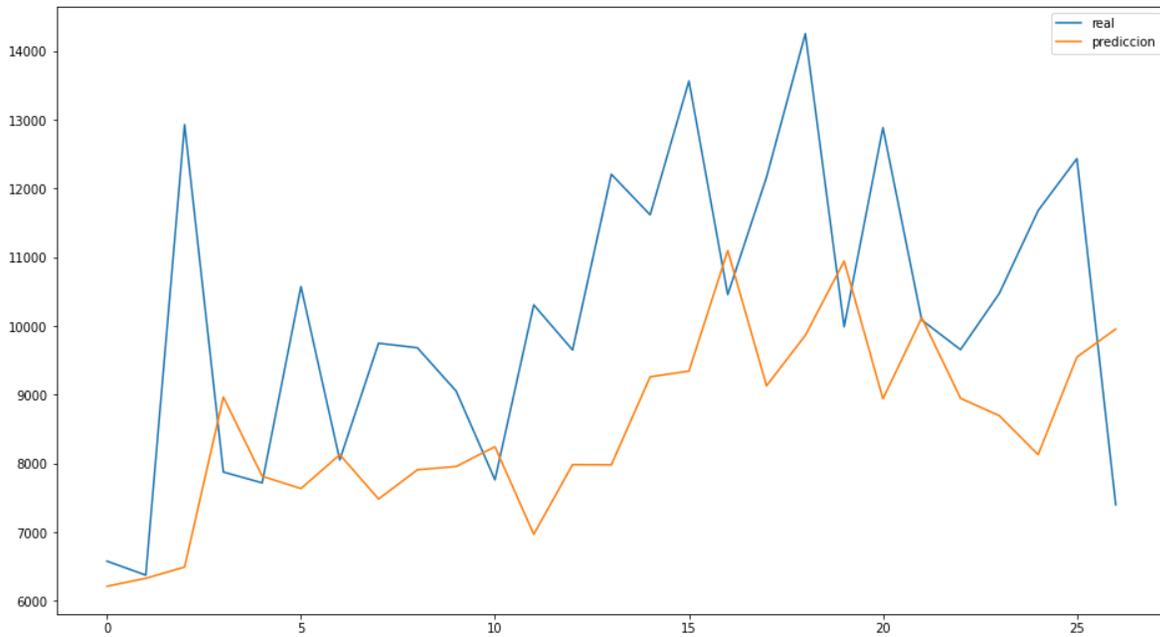
Anexo i. Predicción vs datos reales de la categoría de lomo de cerdo con un mape de 9,19% de error.



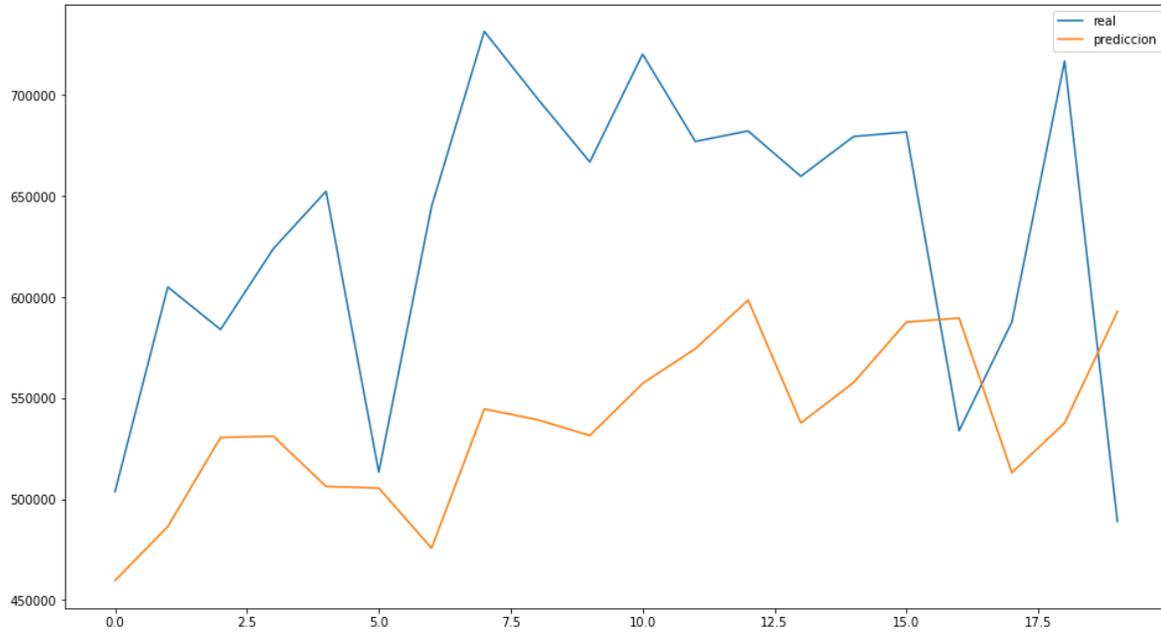
Anexo j. Predicción vs datos reales de la categoría de paleta de cerdo con un mape de 14,34% de error.



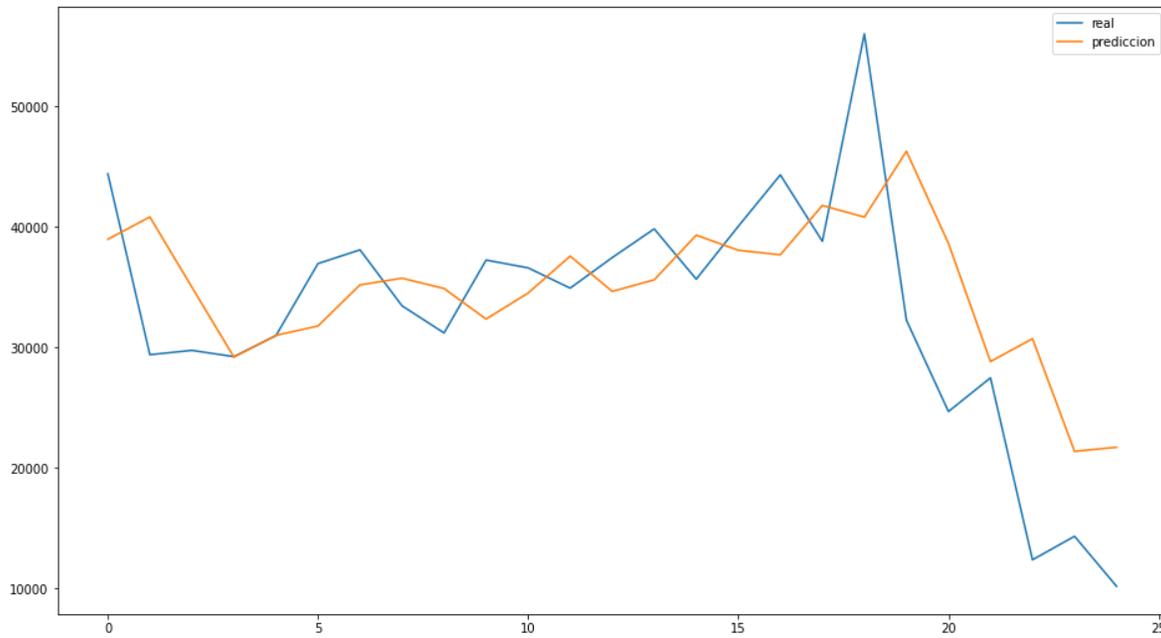
Anexo k. Predicción vs datos reales de la categoría de panceta de cerdo con un mape de 61,46% de error.



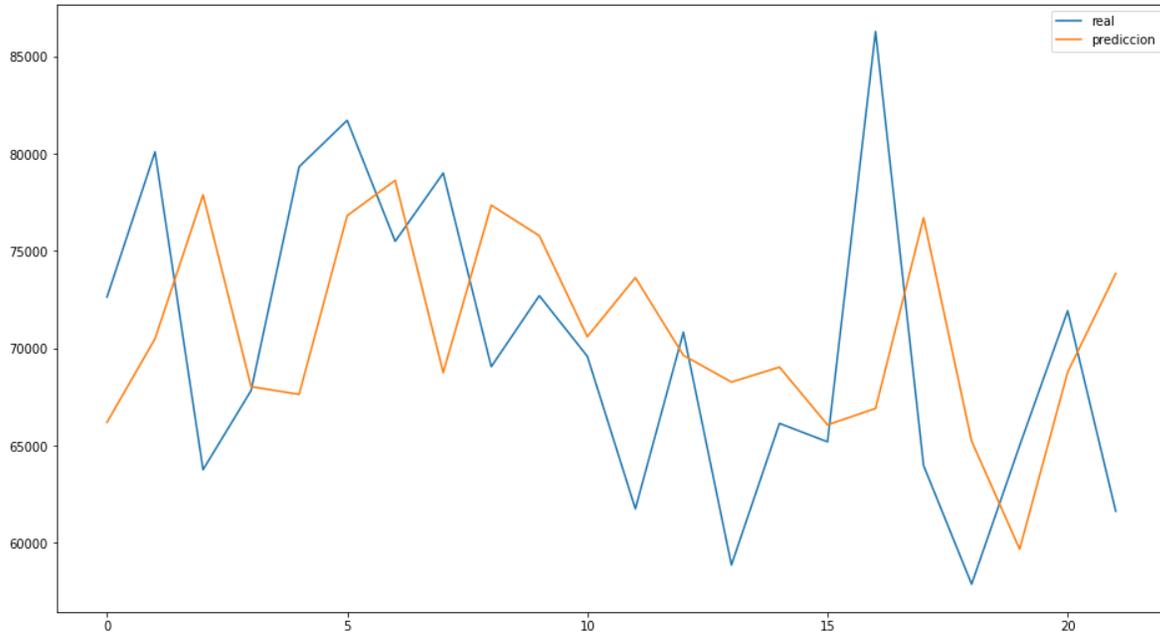
Anexo l. Predicción vs datos reales de la categoría de perril de cerdo con un mape de 18,89% de error.



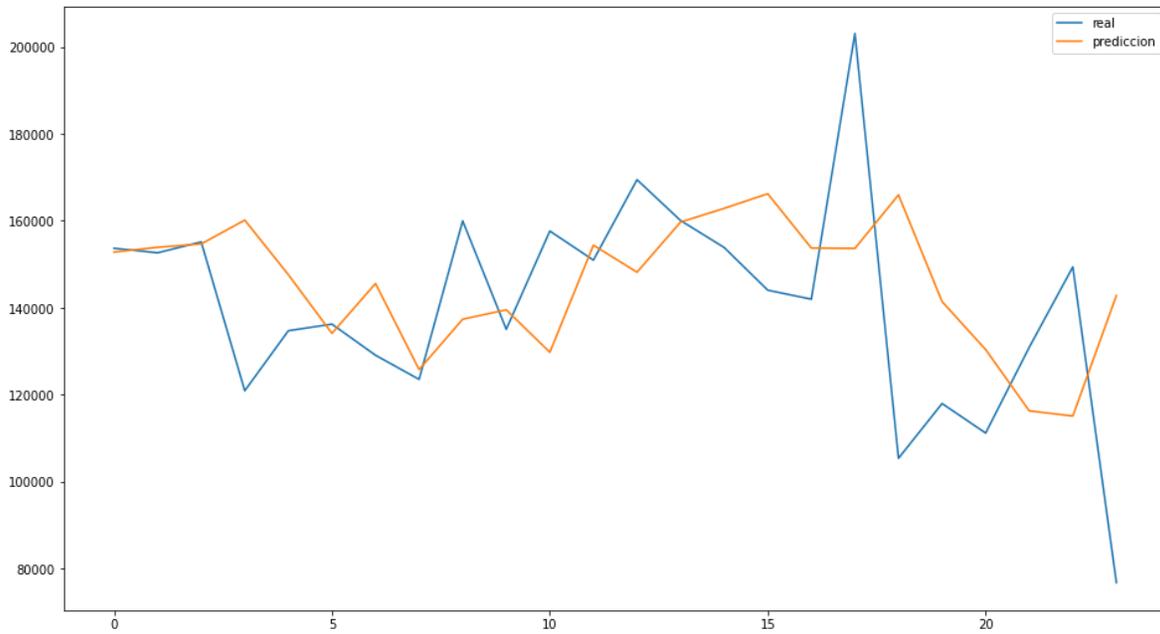
Anexo m. Predicción vs datos reales de la categoría de pierna de cerdo con un mape de 17,04% de error.



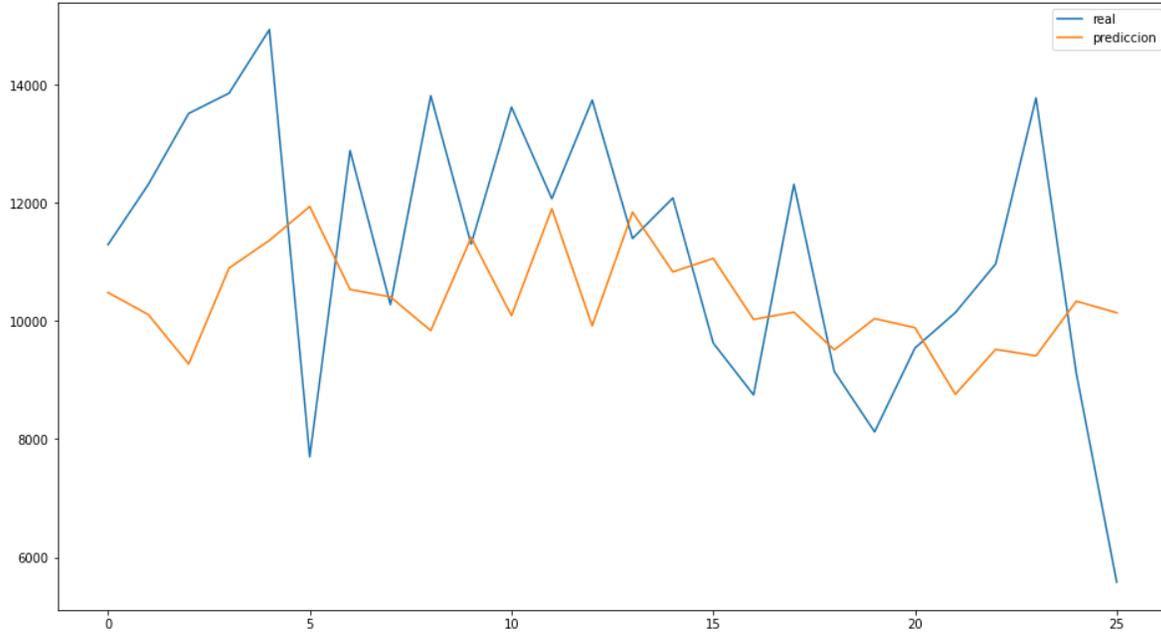
Anexo n. Predicción vs datos reales de la categoría de plancha de cerdo con un mape de 25,33% de error.



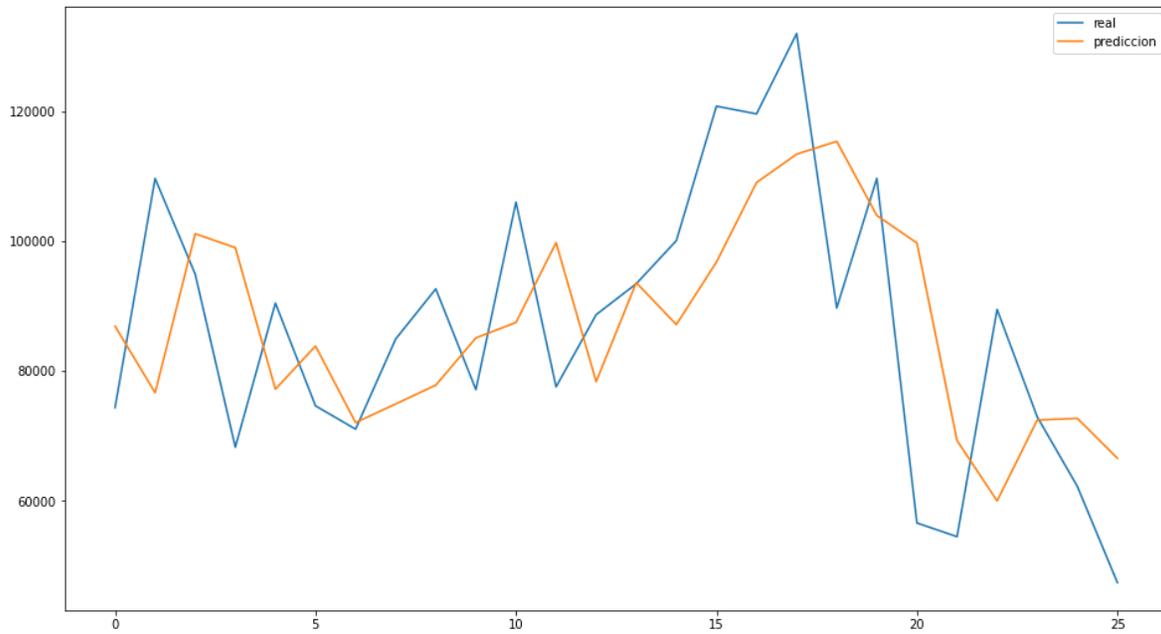
Anexo o. Predicción vs datos reales de la categoría de prolijado de cerdo con un mape de 10,4% de error.



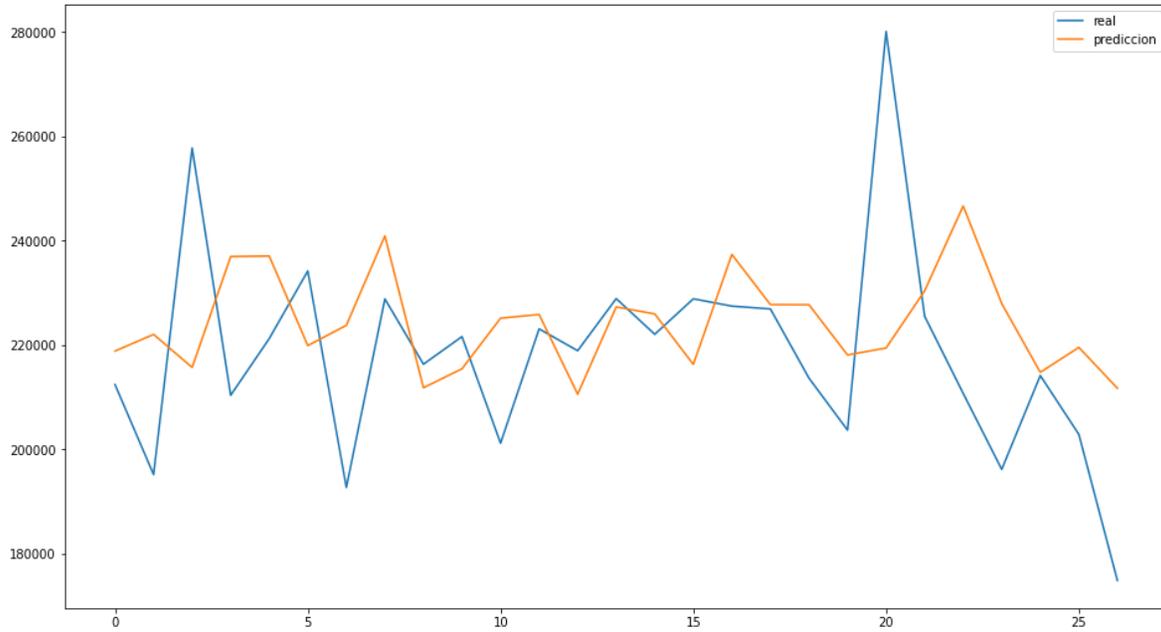
Anexo p. Predicción vs datos reales de la categoría de recortes de cerdo con un mape de 15,77% de error.



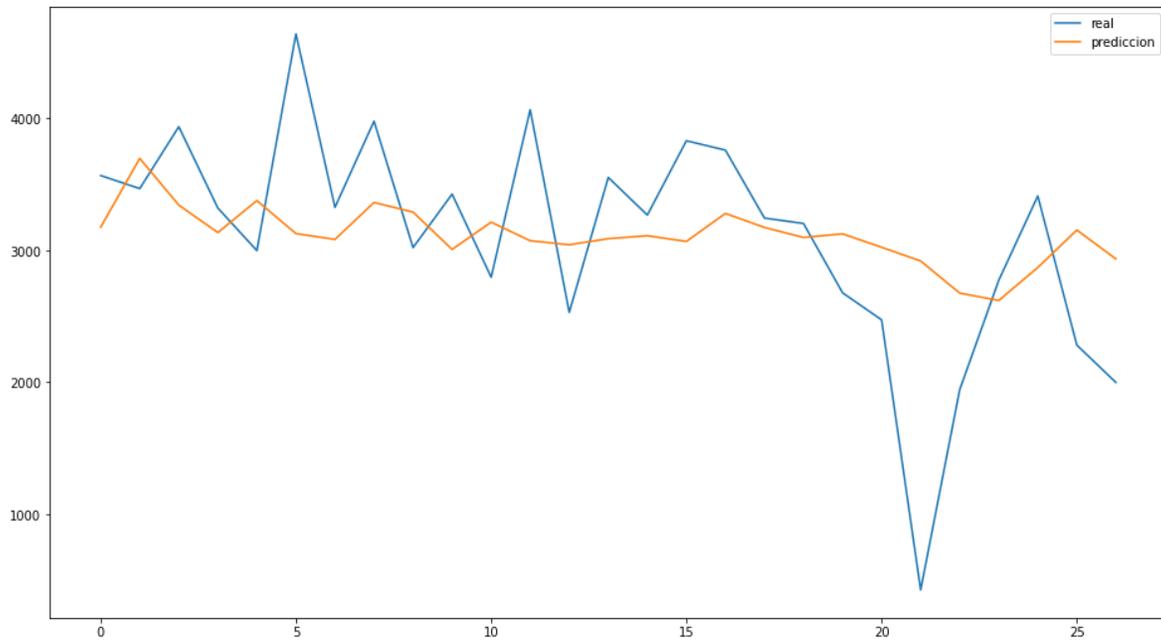
Anexo q. Predicción vs datos reales de la categoría de reproductor de cerdo con un mape de 19,49% de error.



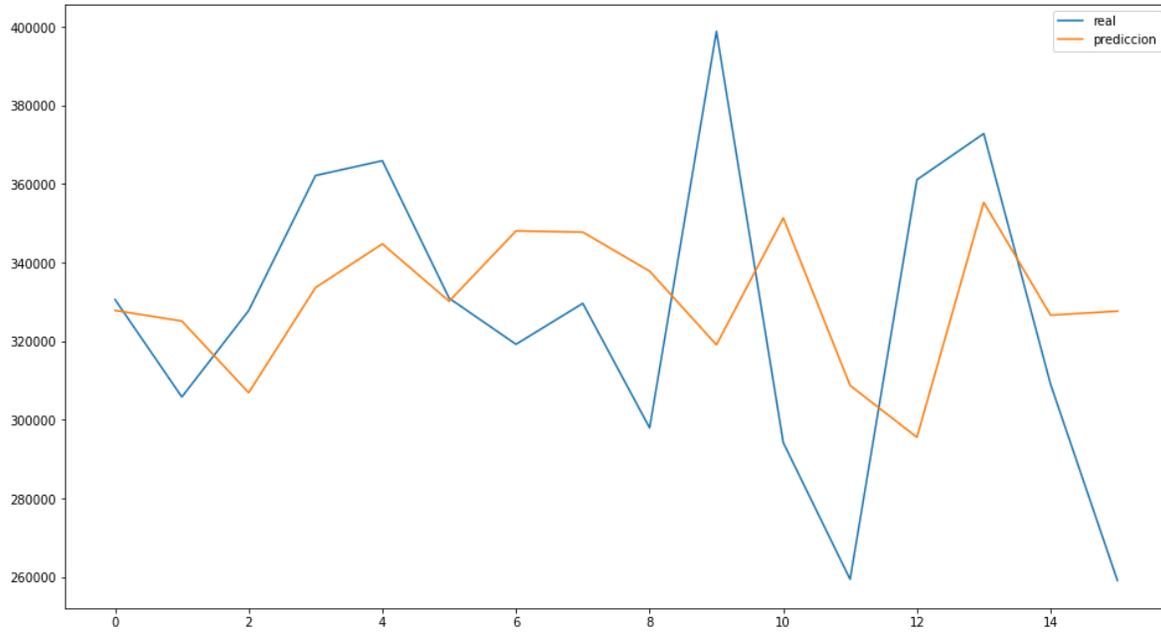
Anexo r. Predicción vs datos reales de la categoría de subproductos de cerdo con un mape de 19,52% de error.



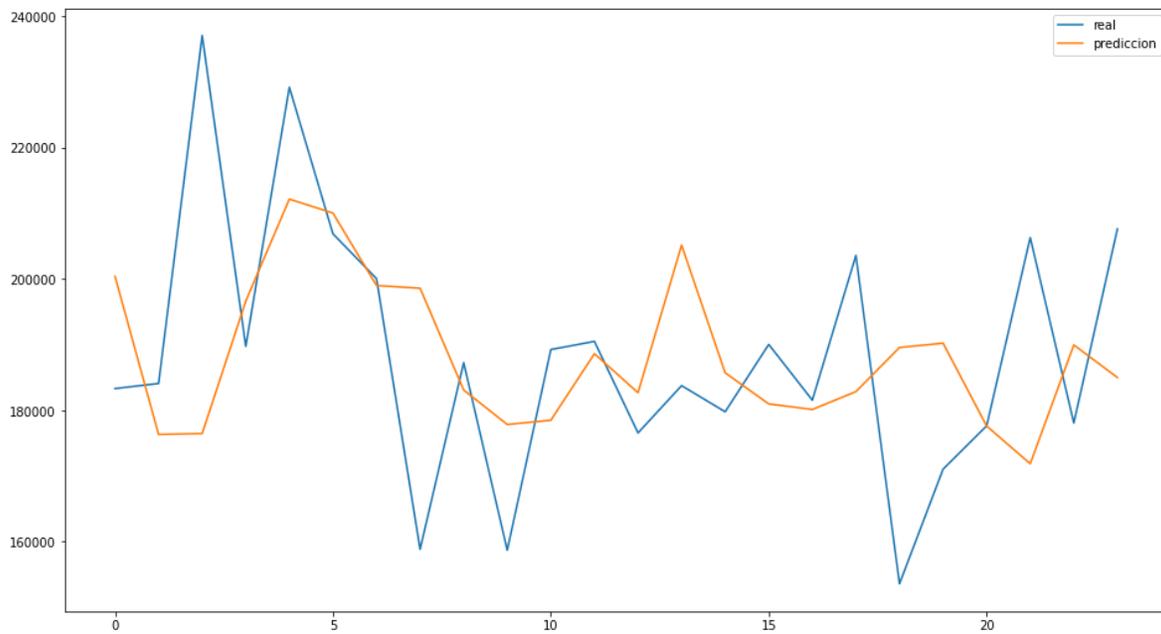
Anexo s. Predicción vs datos reales de la categoría de ala de pollo con un mape de 7,97% de error.



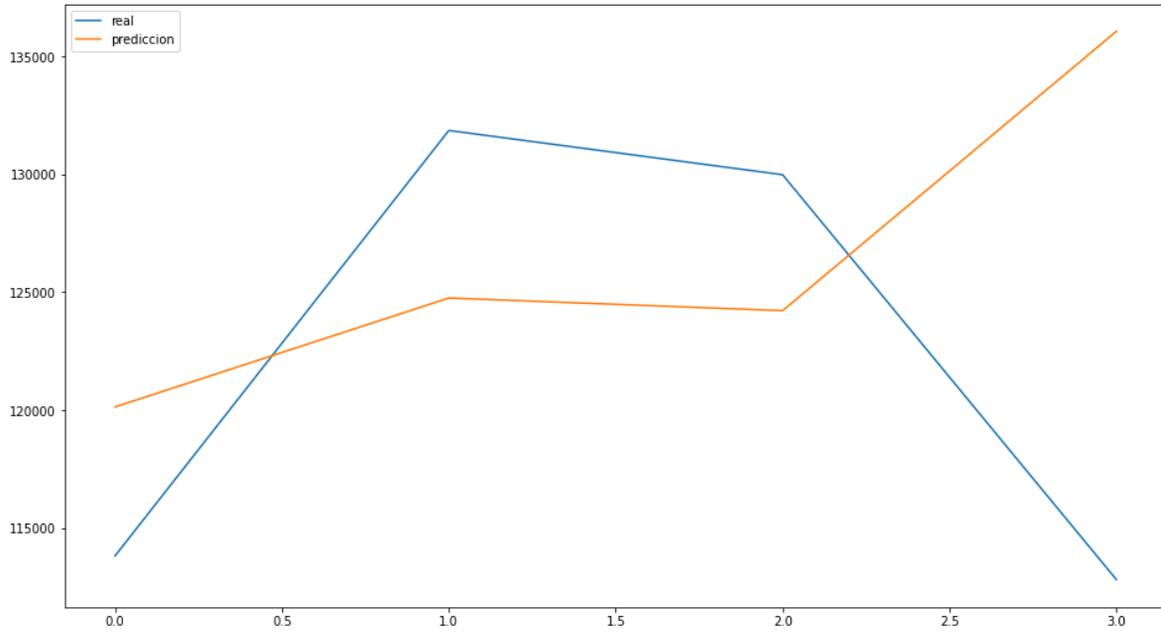
Anexo t. Predicción vs datos reales de la categoría de cazuela de pollo con un mape de 37,46% de error.



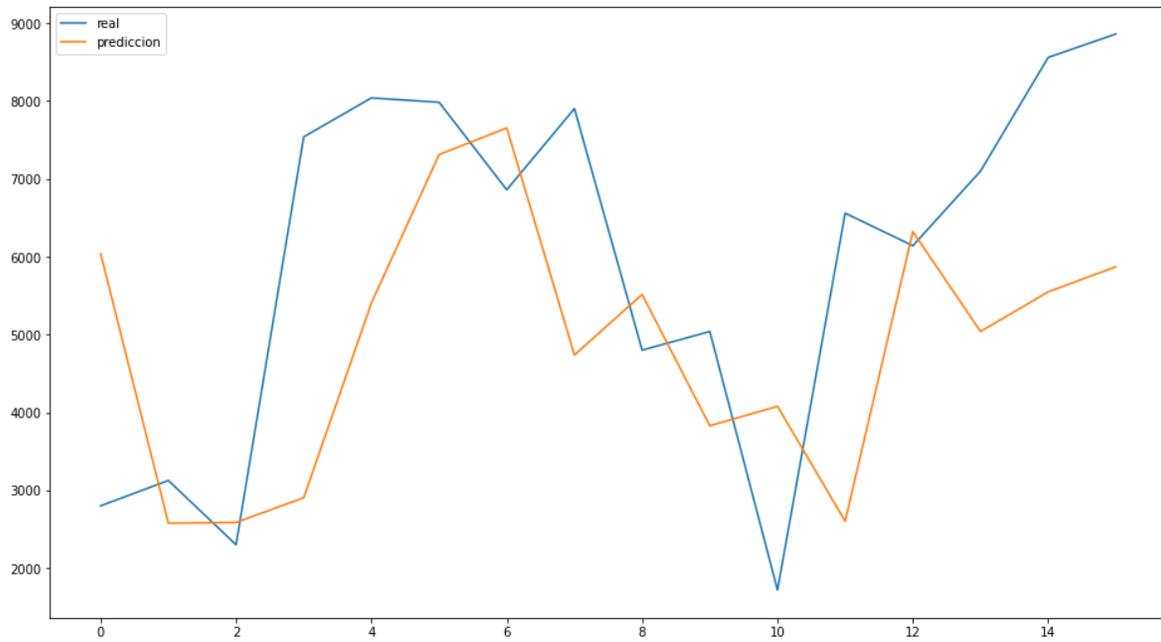
Anexo u. Predicción vs datos reales de la categoría de con menudencias de pollo con un mape de 10,54% de error.



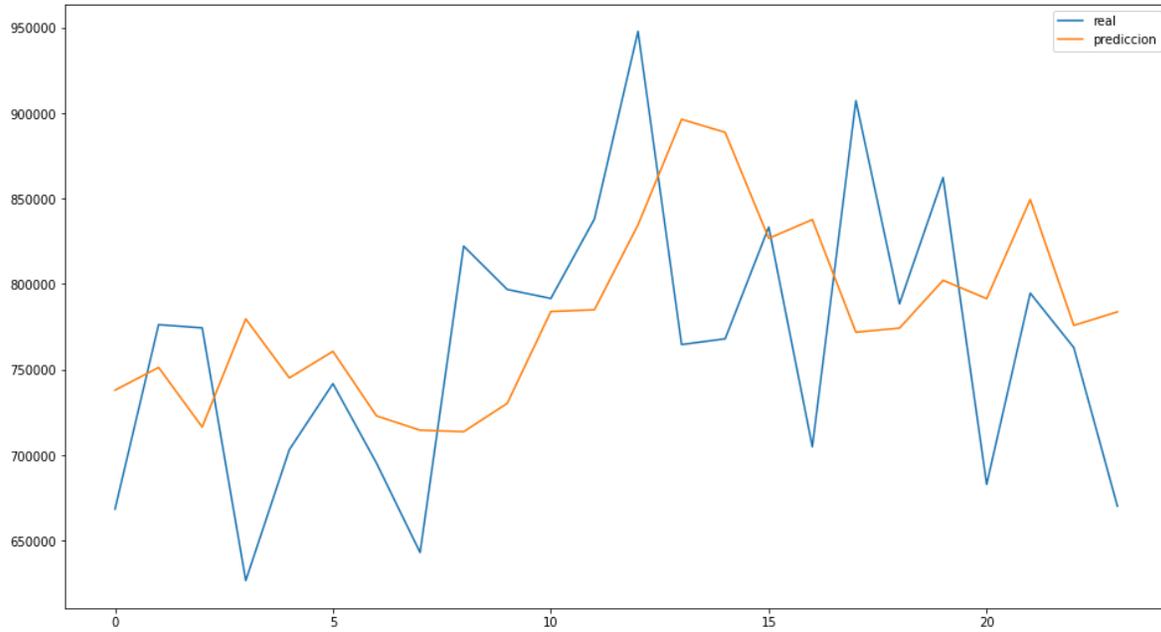
Anexo v. Predicción vs datos reales de la categoría de filete de pollo con un mape de 8,38% de error.



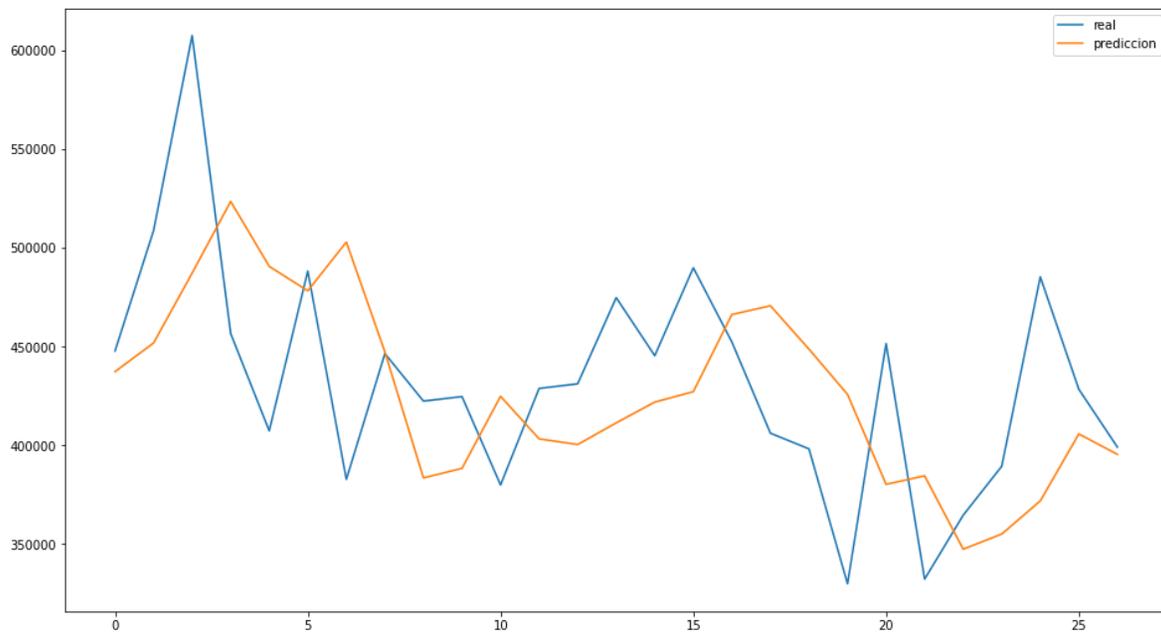
Anexo x. Predicción vs datos reales de la categoría de menudencias de pollo con un mape de 8,99% de error.



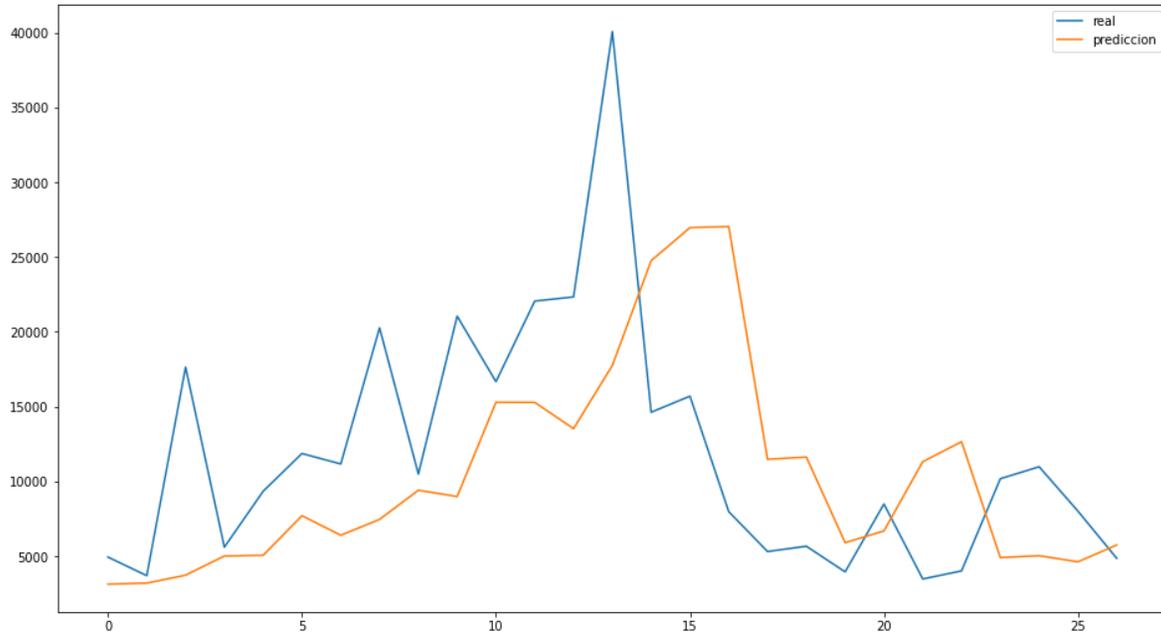
Anexo y. Predicción vs datos reales de la categoría de patas de pollo con un mape de 39,82% de error.



Anexo z. Predicción vs datos reales de la categoría de pechuga de pollo con un mape de 9,48% de error.



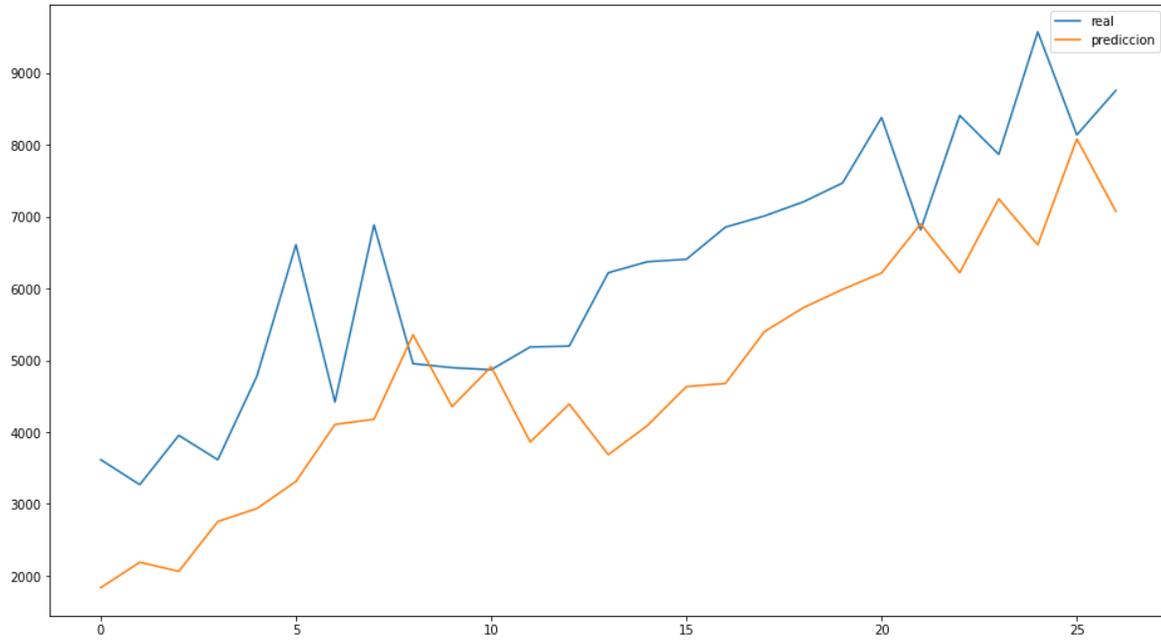
Anexo aa. Predicción vs datos reales de la categoría de pechuga deshuesada de pollo con un mape de 11,53% de error.



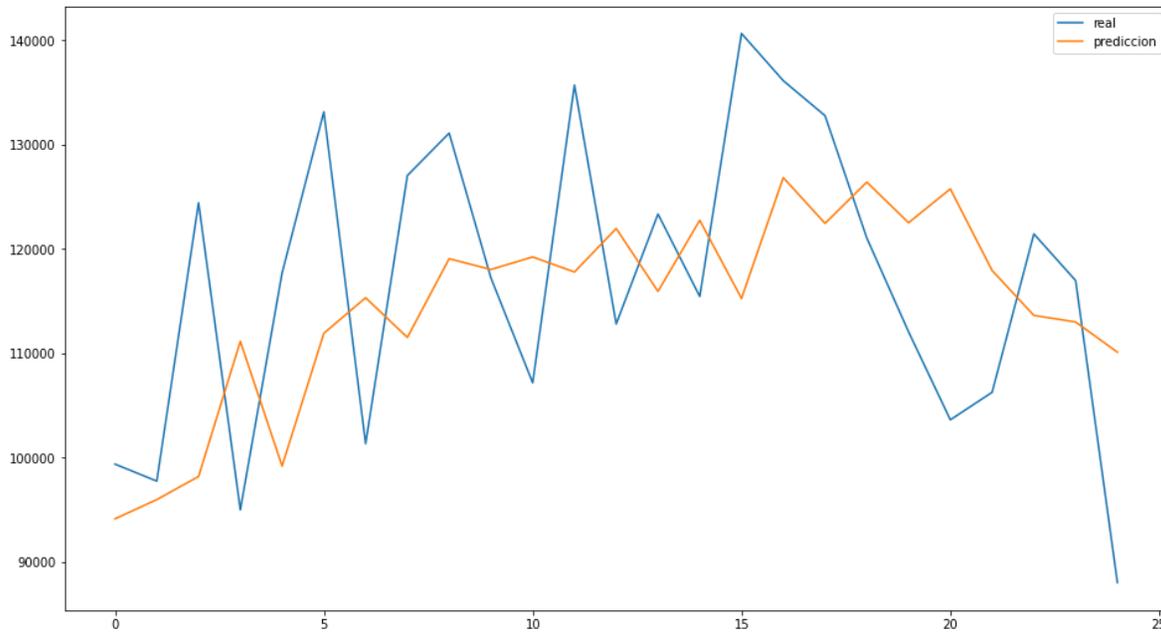
Anexo ab. Predicción vs datos reales de la categoría de reproductor de pollo con un mape de 66,95% de error.



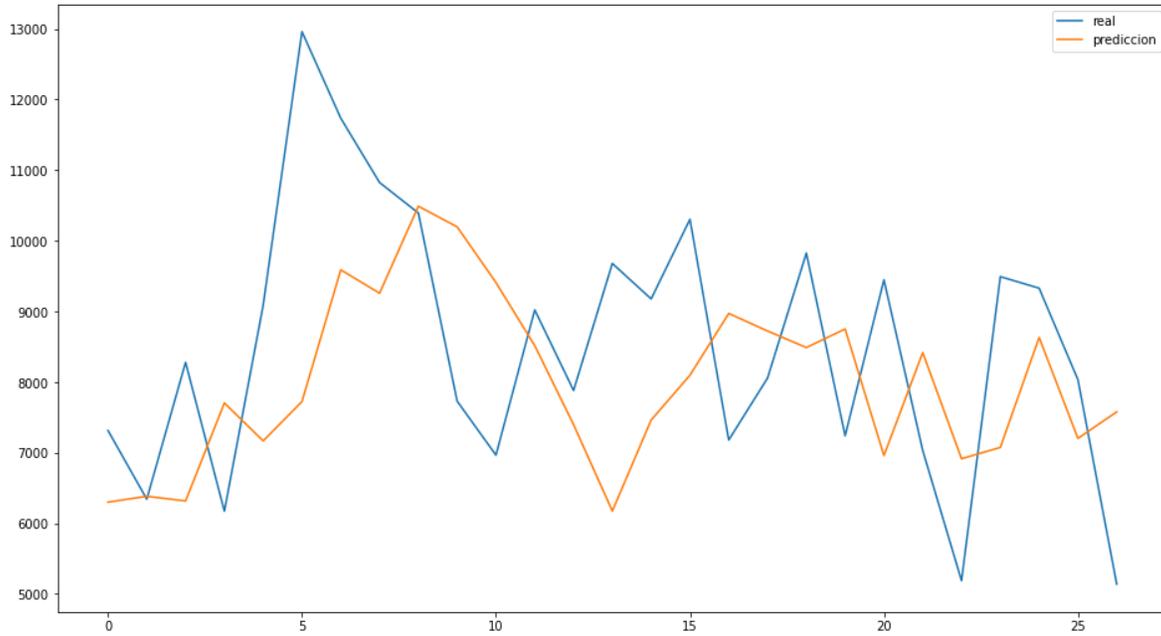
Anexo ac. Predicción vs datos reales de la categoría de trutro de pollo con un mape de 7,57% de error.



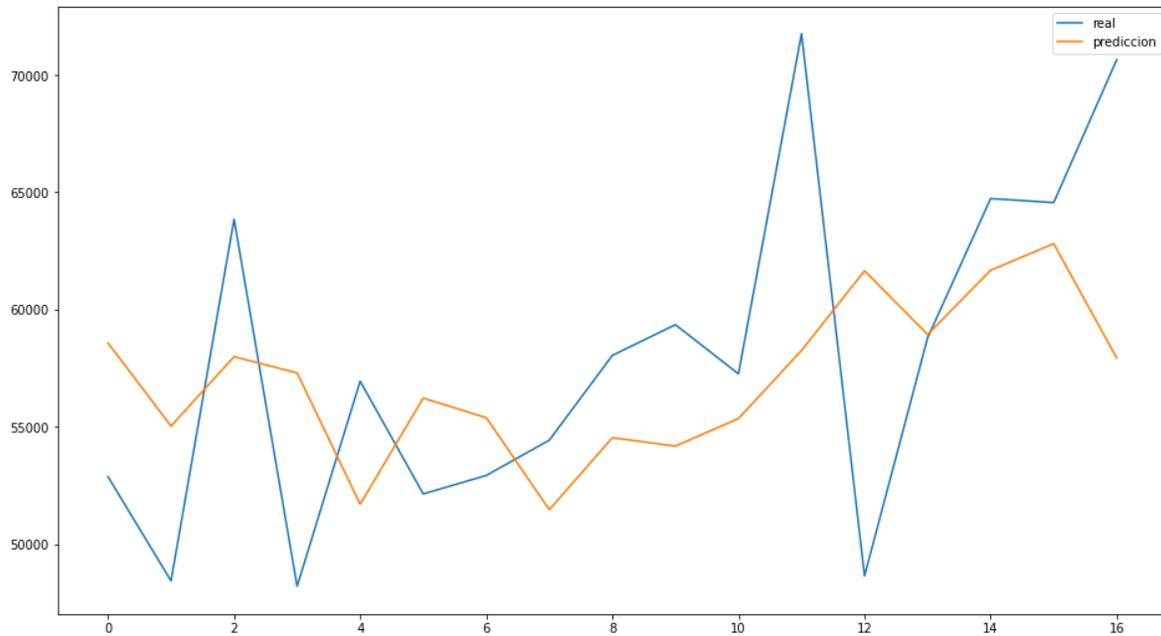
Anexo ad. Predicción vs datos reales de la categoría de trutro deshuesado de pollo con un mape de 24,48% de error.



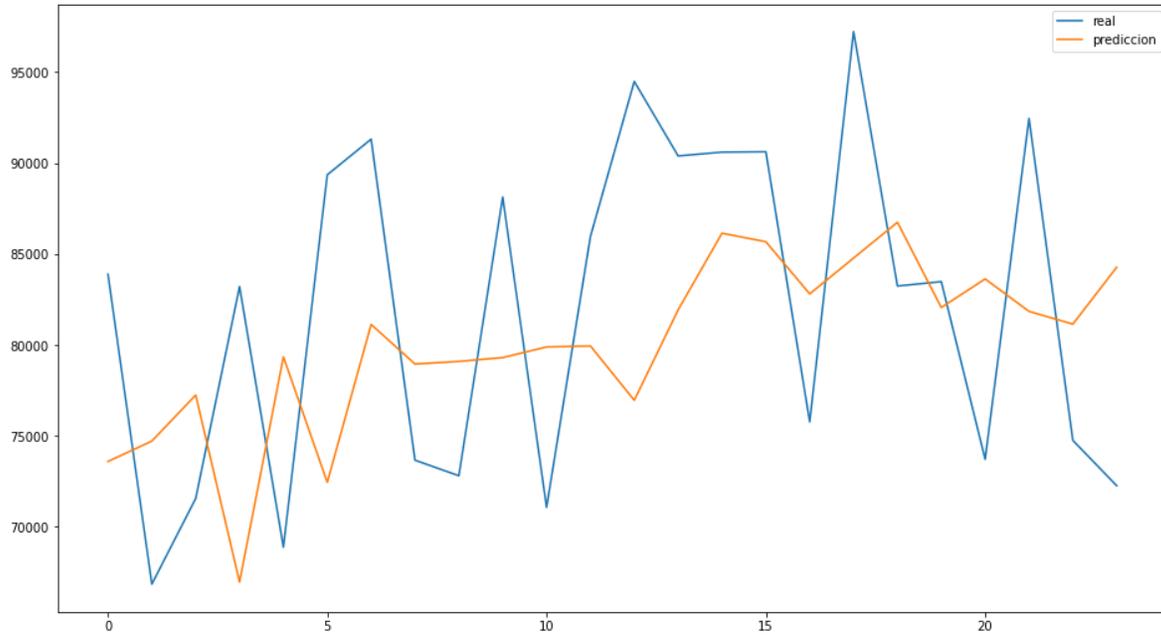
Anexo ae. Predicción vs datos reales de la categoría de ala de pavo con un mape de 10,86% de error.



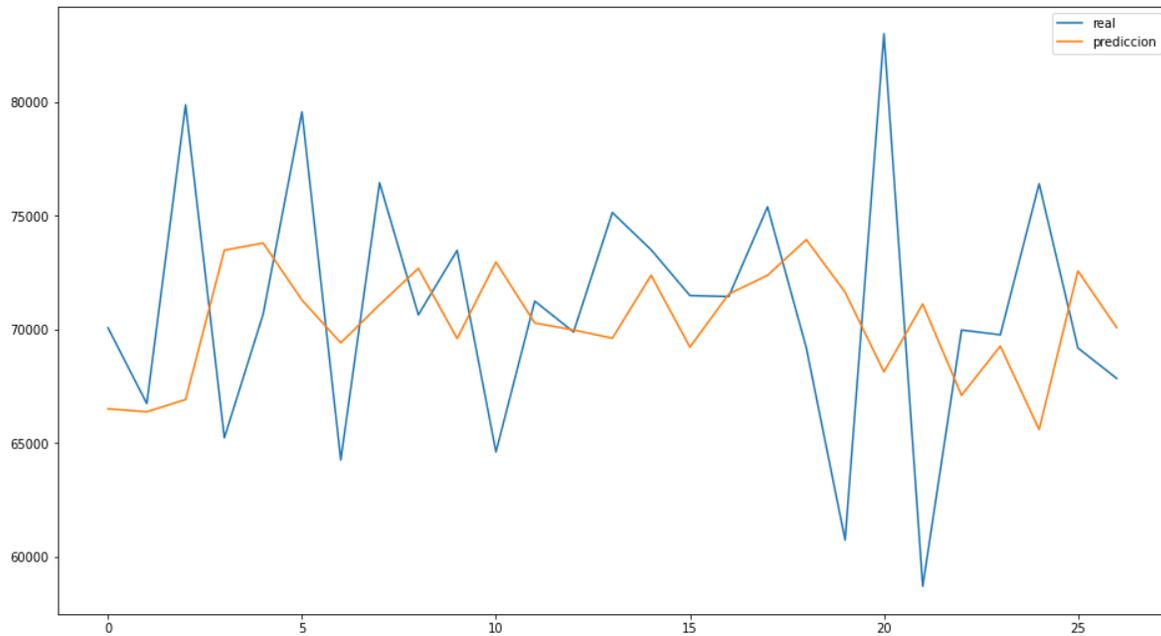
Anexo af. Predicción vs datos reales de la categoría de menudencias de pavo con un mape de 20,41% de error.



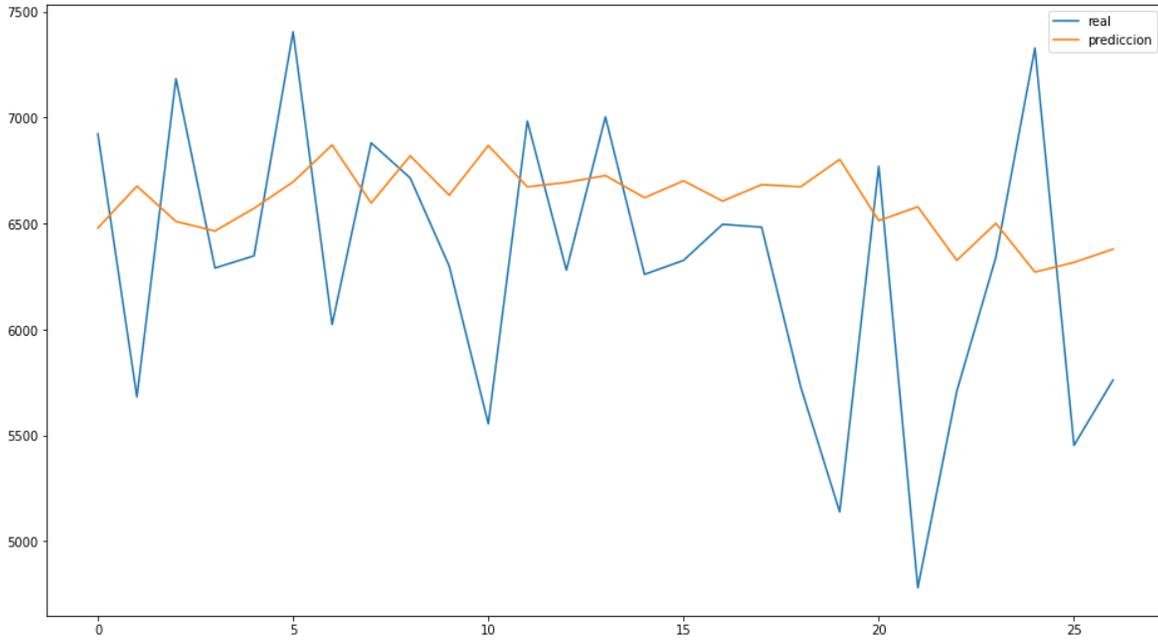
Anexo ag. Predicción vs datos reales de la categoría de pechuga deshuesada de pavo con un mape de 9,93% de error.



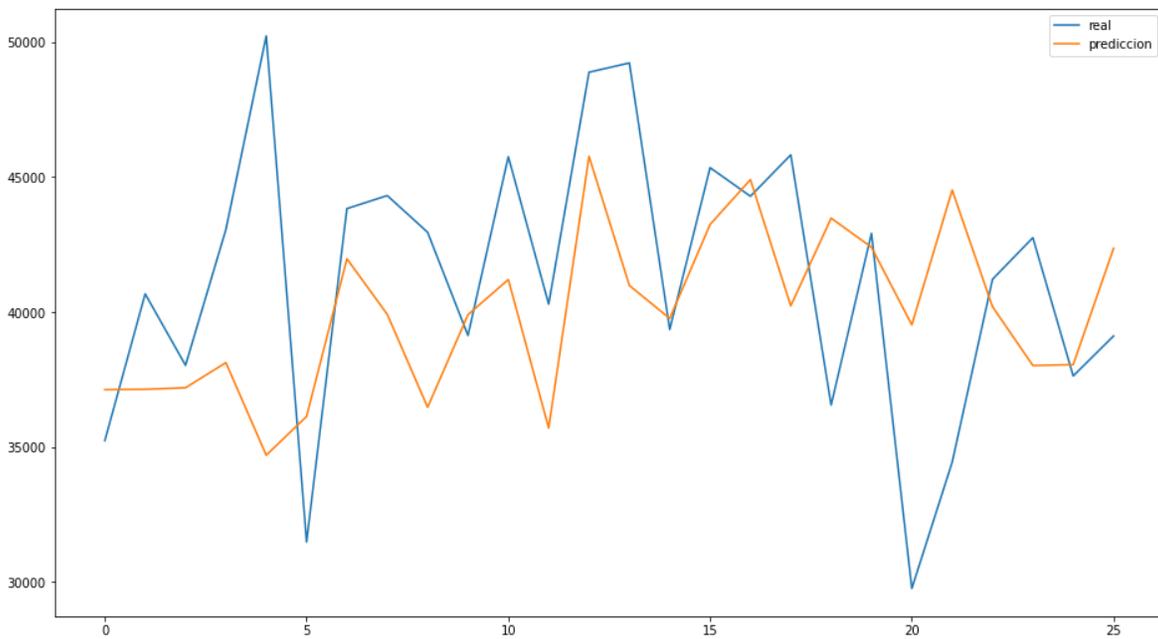
Anexo ah. Predicción vs datos reales de la categoría de trutro de pavo con un mape de 10,76% de error.



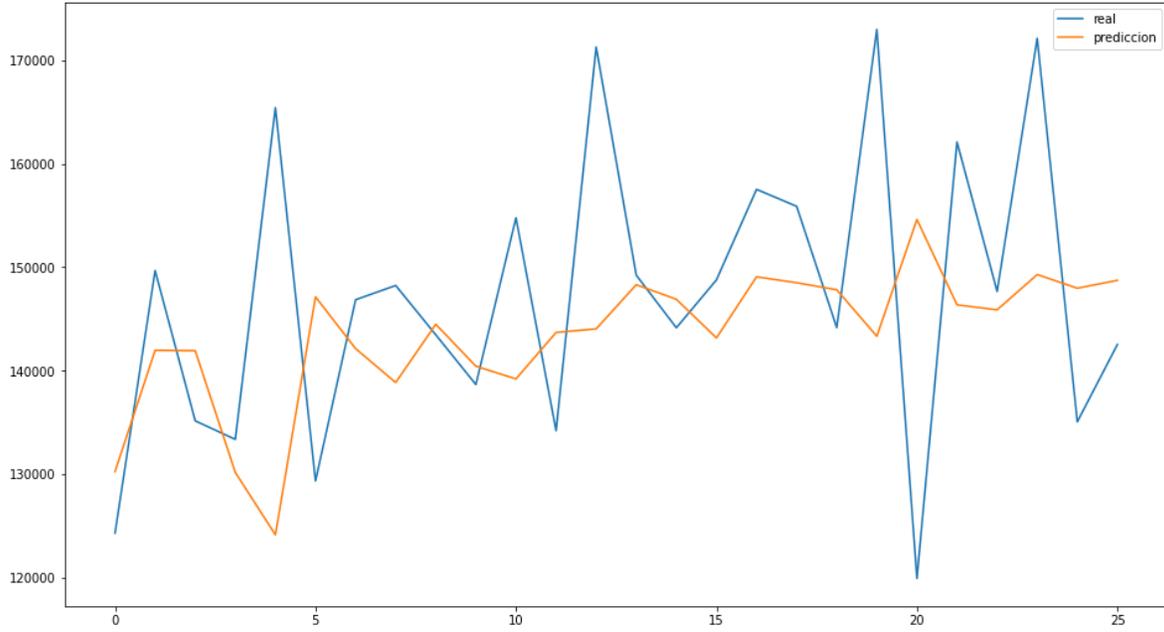
Anexo ai. Predicción vs datos reales de la categoría de trutro deshuesado de pavo con un mape de 7,19% de error.



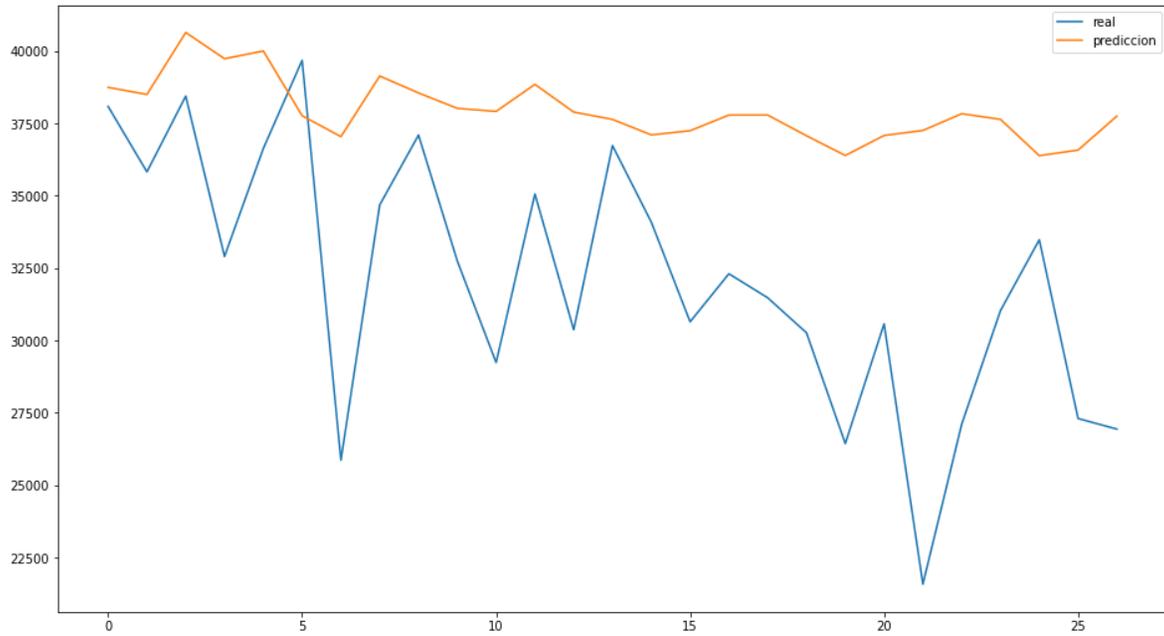
Anexo aj. Predicción vs datos reales de la categoría de arrollado de cecina con un mape de 10,14% de error.



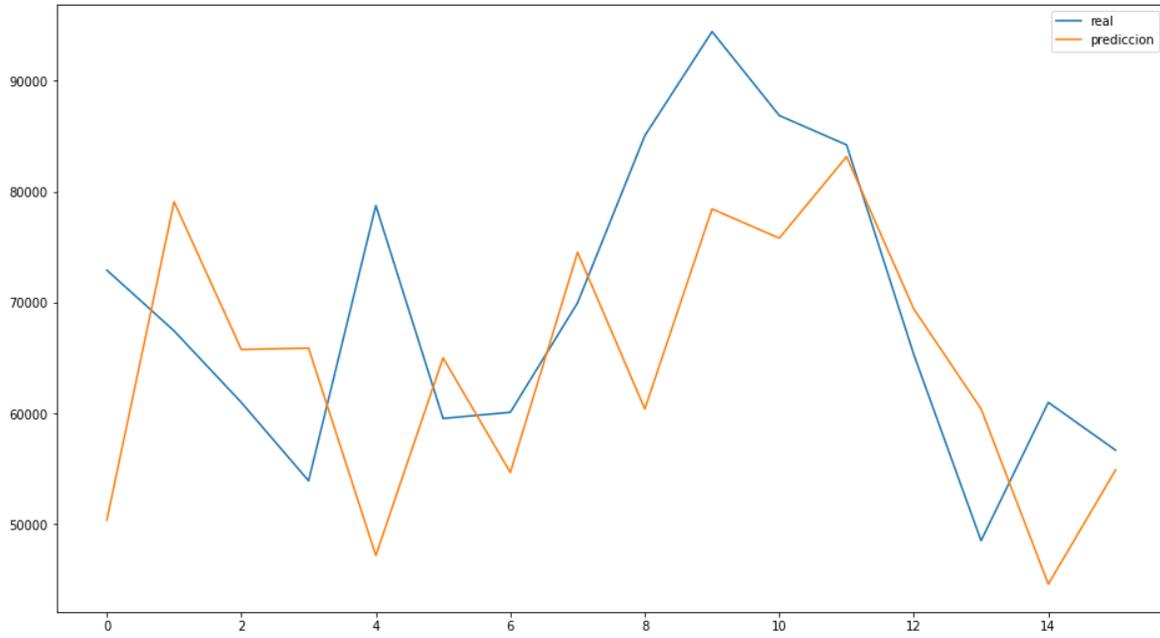
Anexo ak. Predicción vs datos reales de la categoría de fiambre de cecina con un mape de 10,52% de error.



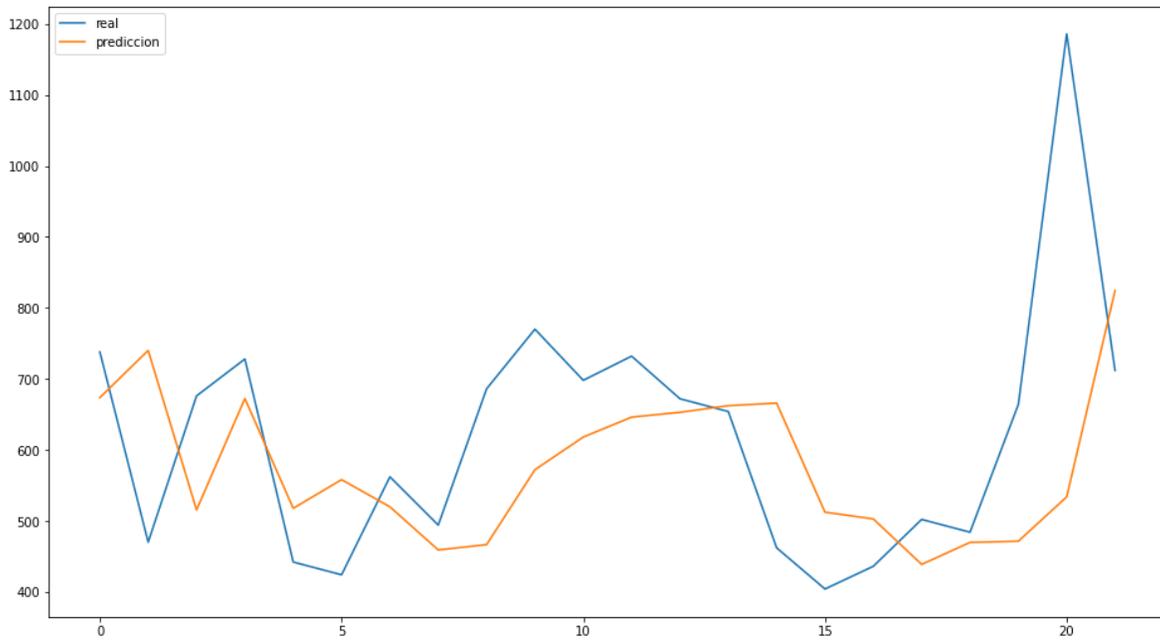
Anexo al. Predicción vs datos reales de la categoría de jamón de cecina con un mape de 7,84% de error.



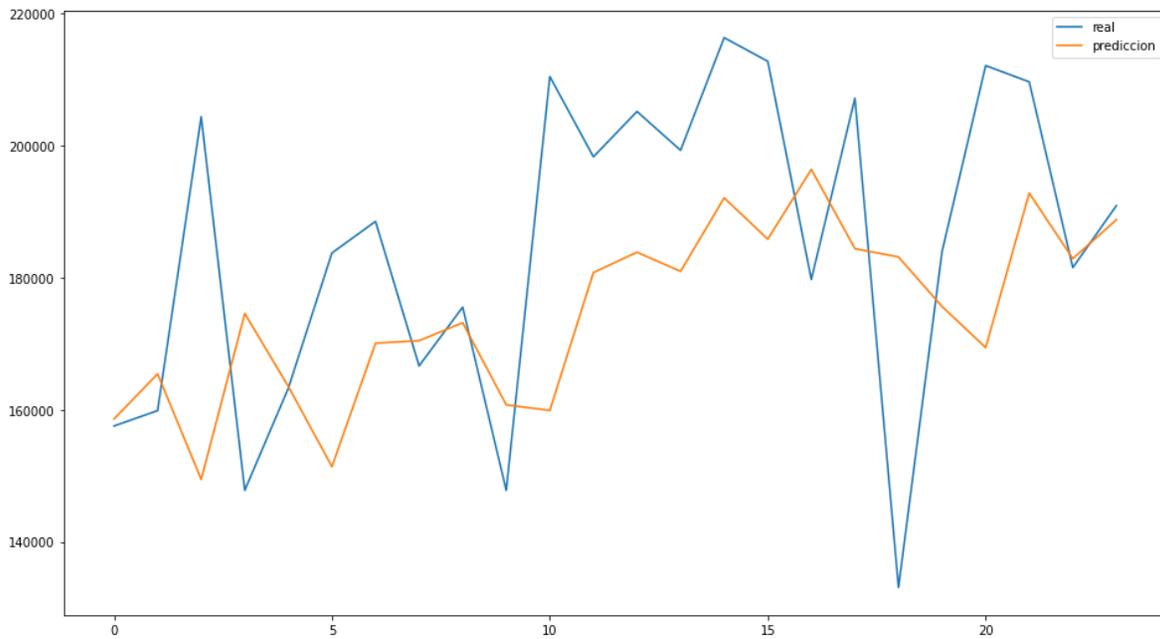
Anexo am. Predicción vs datos reales de la categoría de mortadela de cecina con un mape de 20,7% de error.



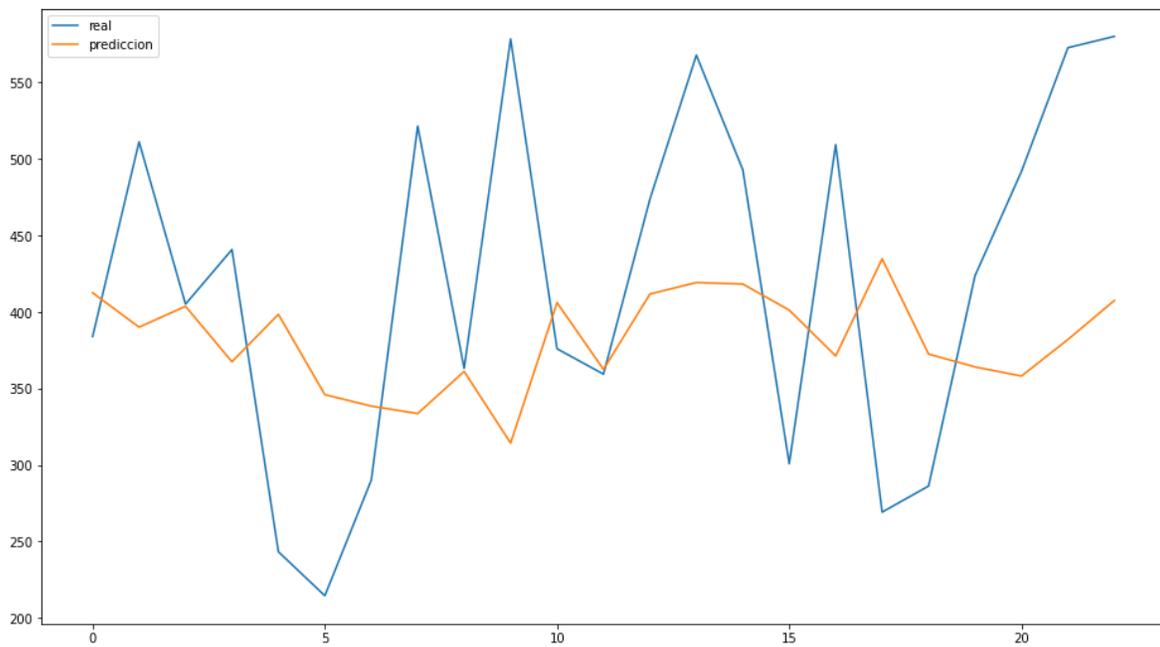
Anexo an. Predicción vs datos reales de la categoría de parrilleros de cecina con un mape de 16,5% de error.



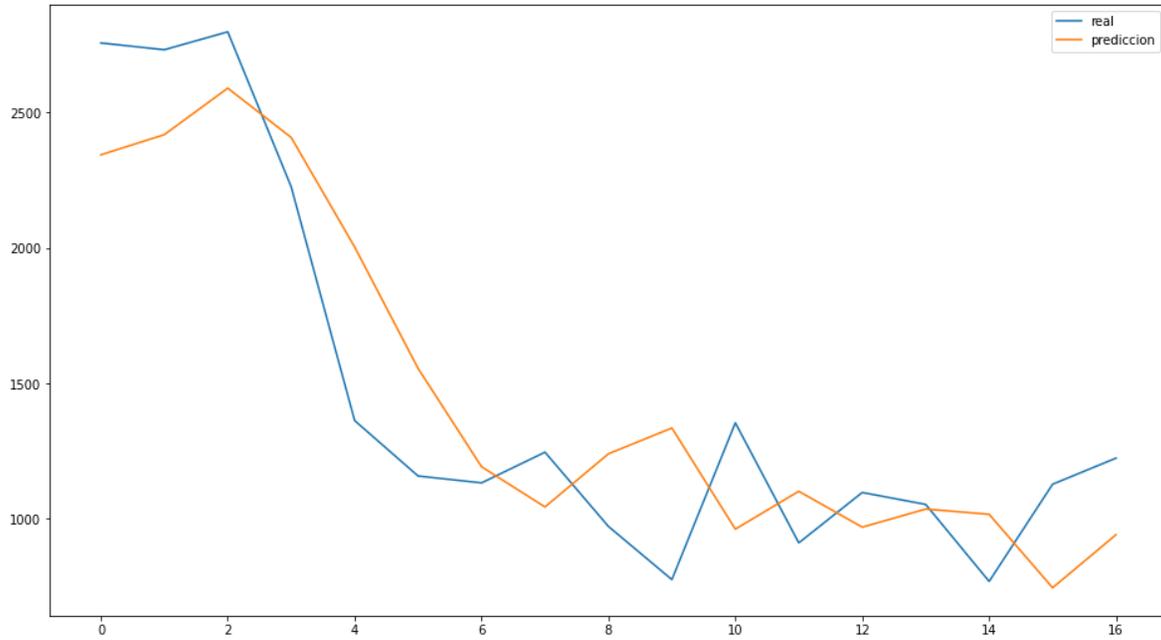
Anexo ao. Predicción vs datos reales de la categoría de paté de cecina con un mape de 20,34% de error.



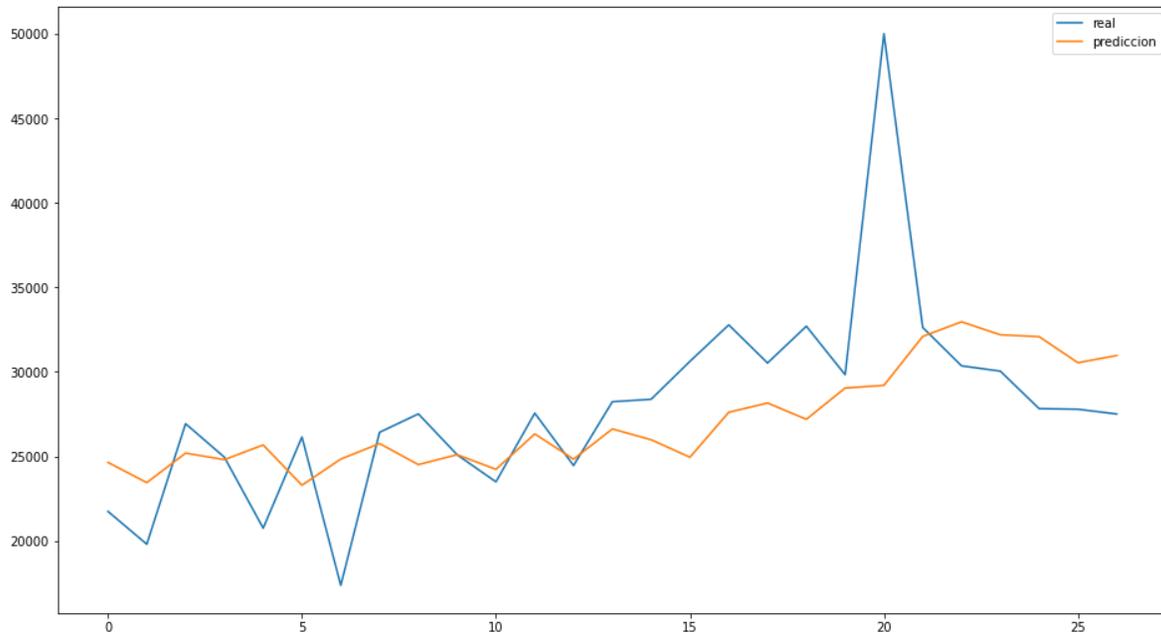
Anexo ap. Predicción vs datos reales de la categoría de pechuga de cecina con un mape de 10,74% de error.



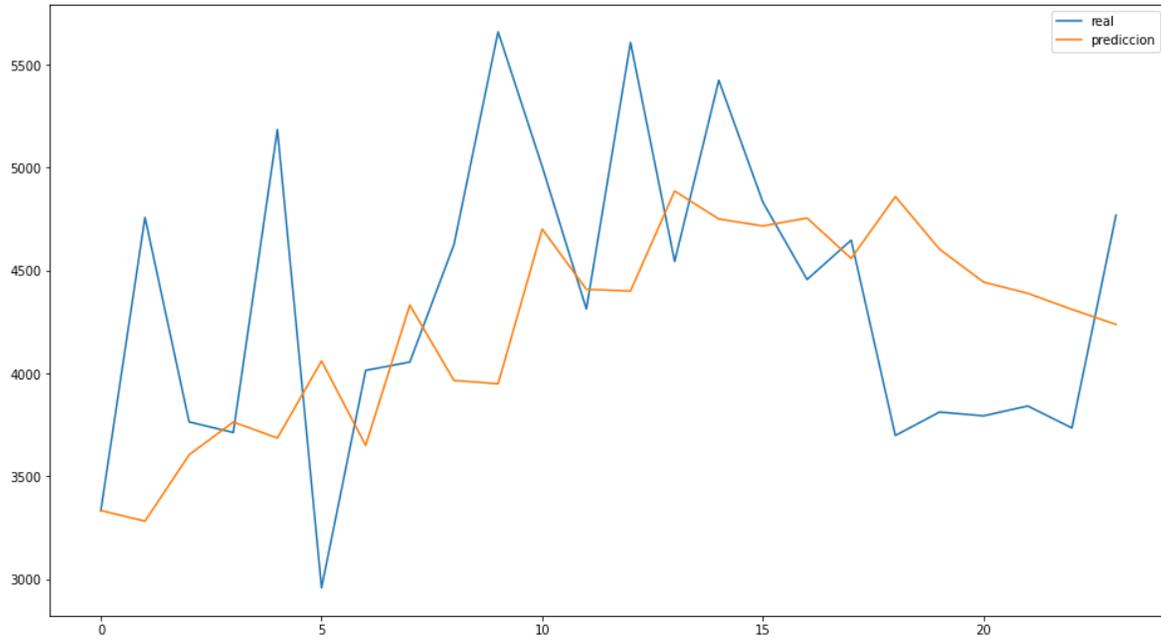
Anexo aq. Predicción vs datos reales de la categoría de productos especiales de cecina con un mape de 25,72% de error.



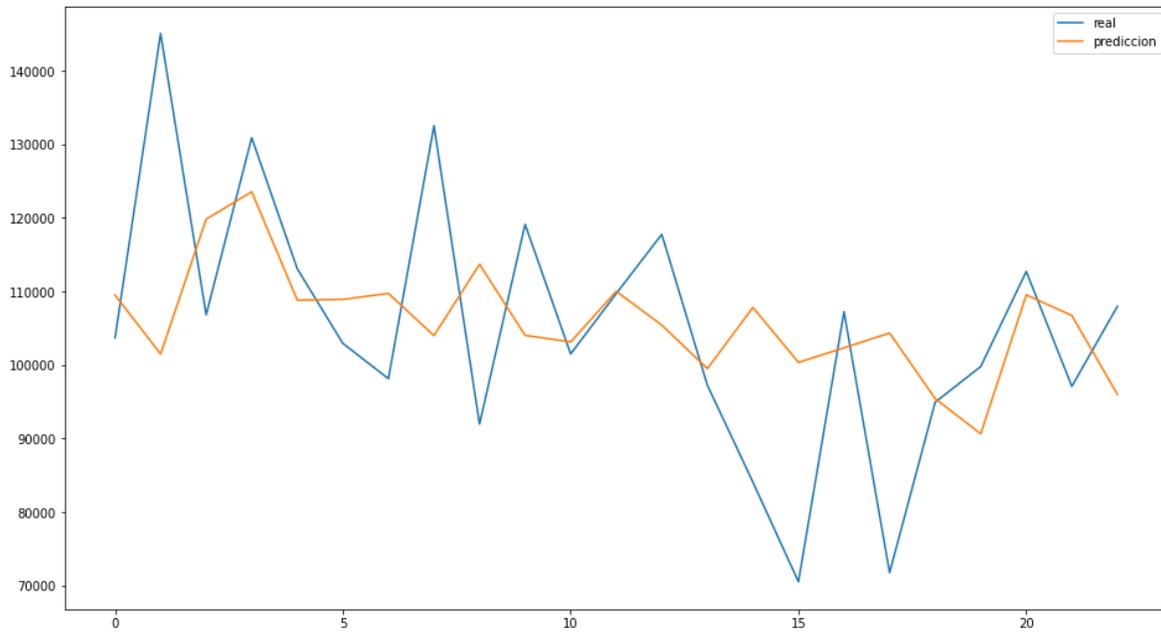
Anexo ar. Predicción vs datos reales de la categoría de salame de cecina con un mape de 23,34% de error.



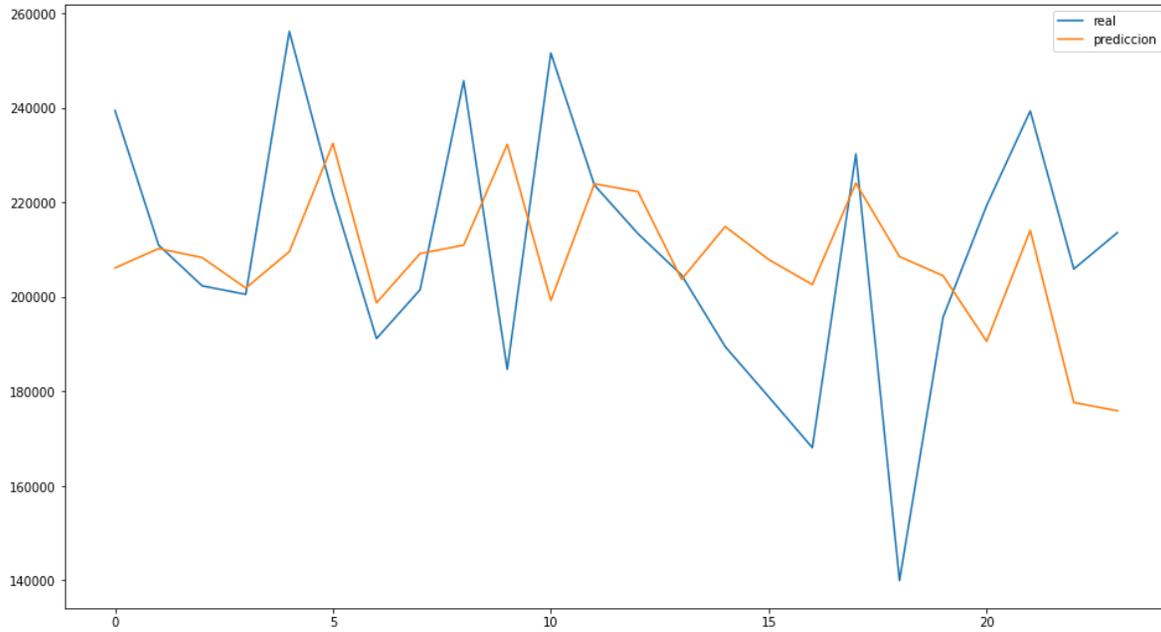
Anexo as. Predicción vs datos reales de la categoría de salchicha de cecina con un mape de 11,52% de error.



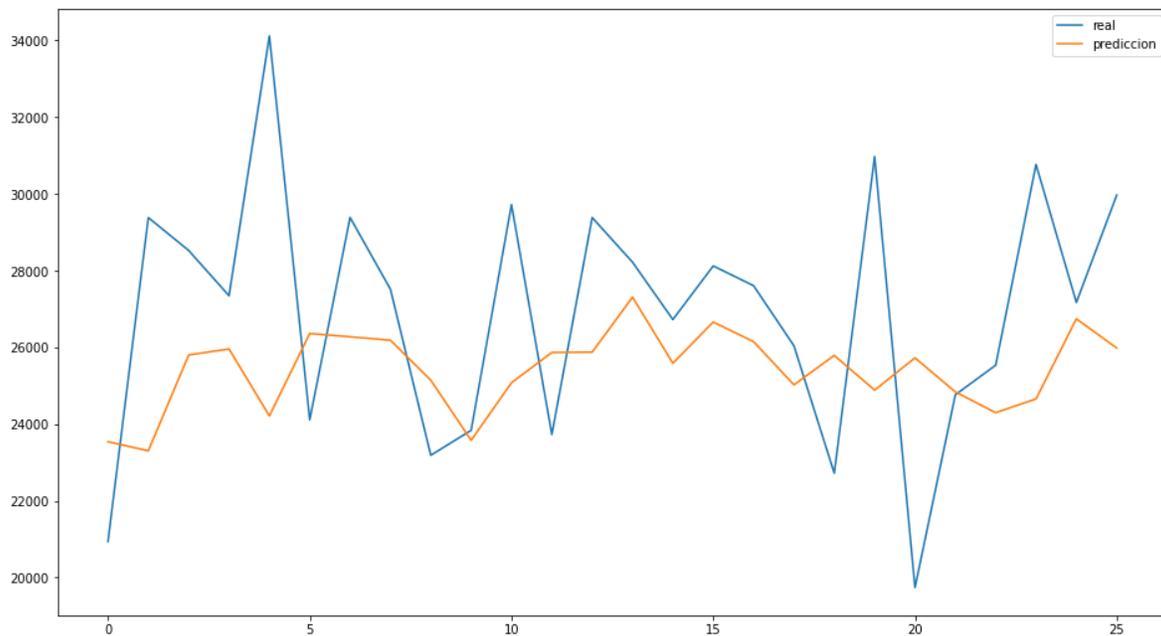
Anexo at. Predicción vs datos reales de la categoría de empanadas de elaborado con un mape de 13,91% de error.



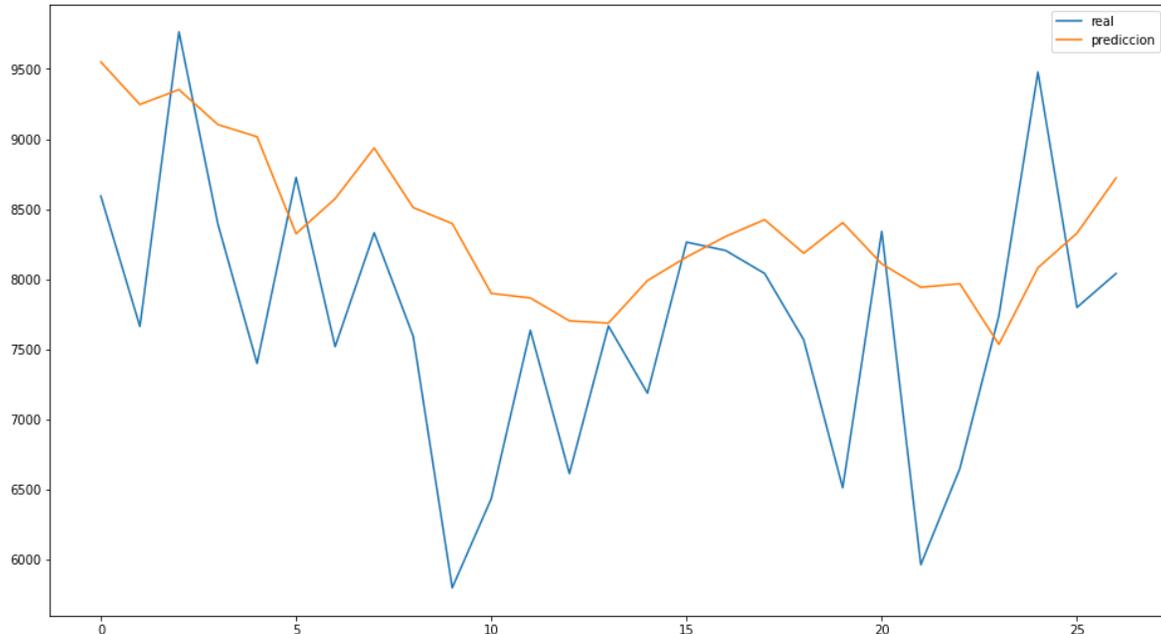
Anexo au. Predicción vs datos reales de la categoría de empanizado de elaborado con un mape de 13,09% de error.



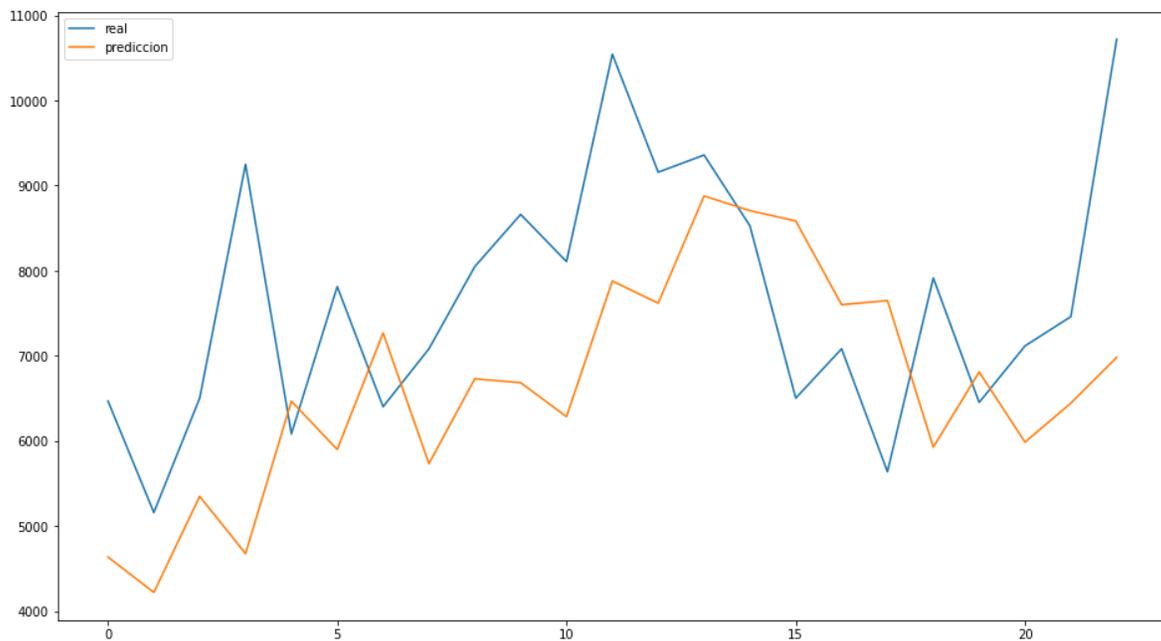
Anexo av. Predicción vs datos reales de la categoría de hamburguesa de elaborado con un mape de 11,48% de error.



Anexo aw. Predicción vs datos reales de la categoría de moldeado de elaborado con un mape de 10,49% de error.



Anexo ax. Predicción vs datos reales de la categoría de molido de elaborado con un mape de 12,44% de error.



Anexo ay. Predicción vs datos reales de la categoría de molido de elaborado con un mape de 19,92% de error.