

Proyecto: "Forecast de ventas de repuestos automotrices"

Actividad Final de Graduación III - Magíster en Ciencias de Datos UC

Informe Final

Grupo Elena16

Integrantes:

- Bastián Gutiérrez
- Sebastián Gamboa
- Luciano Davico

Resumen

La gestión eficiente del inventario de repuestos automotrices es un desafío crucial en la industria automotriz. La capacidad para predecir con precisión las ventas futuras de repuestos es esencial para equilibrar la disponibilidad de productos y los costos asociados, evitando tanto quiebres de stock, caso en el cual se pierde venta, como también en excesos de inventario que conlleva altos gastos de almacenamiento. Este proyecto se centra en el estudio de modelos predictivos para anticipar la demanda de tipo intermitente de repuestos automotrices, mediante el lenguaje de programación Python y modelos de machine learning y análisis de series de tiempo, en donde se prioriza el modelo de Croston, el cual logra las mejores predicciones para este caso de uso. El estudio se desarrolló usando la metodología CRISP-DM y contemplando aspectos éticos importantes en la implementación de los modelos.

1. Introducción

La industria automotriz actualmente enfrenta como uno de sus principales retos la gestión eficiente de inventario de repuestos, es importante que esto se gestione de la manera más eficiente posible, ya que un mal manejo puede llevar a:

- **Quiebres de stock:** Lo cual corresponde a tener una cantidad de inventario incapaz de cubrir la demanda de los clientes, generando pérdida de potenciales ventas, pero por sobre todo, una insatisfacción en los clientes, lo que puede repercutir en pérdidas de venta a mediano y largo plazo.
- **Exceso de stock:** Lo que se refiere a contar con más stock de lo necesario para cubrir la demanda, generando gastos pasivos elevados, riesgo de obsolescencia de los materiales almacenados, en cuyo caso la empresa se ve en la obligación de desechar las piezas que no serán utilizadas en algunos casos.

Es por esto que el uso de técnicas de predicción como lo es machine learning, series de tiempo y deep learning, en conjunto con campos de estudio como la ciencia de datos, son indispensables para apaciguar estas problemáticas que cada vez son más demandadas en el mundo laboral. Con ello, el presente trabajo explora una proyección de la demanda de repuestos que busca:

1.1 Mejorar gestión de inventario, reduciendo así mismo las posibles pérdidas u obsolescencias de productos por sobrestock o quiebres de este mismo esto a través de un Forecast que se explorará, pudiendo influir directamente en la utilidad de la empresa.

1.2 Mantener consideraciones éticas, por este tipo de soluciones siempre está la problemática que afectan el empleo al automatizar soluciones, es por ello que se aborda el tema de la experticia humana como input también importante para dar mayor valor al modelo.

En las secciones siguientes, se describe detalladamente la definición del problema, los objetivos de la investigación, la revisión del estado del arte, el resumen técnico del proyecto con la metodología que se implementó, los resultados que se obtuvieron, la discusión de estos mismos y finalmente las conclusiones, los alcances éticos y las perspectivas de trabajo futuro.

2. Definición del Problema y Objetivos de Investigación

2.1 Definición del Problema

El problema fundamental abordado en este proyecto consiste en **pronosticar la demanda de repuestos automotrices** en un cierto periodo de tiempo determinado, a modo de:

- **Reducir riesgos** asociados a la escasez de productos (quiebre de stock).
- **Optimizar costos** derivados del exceso de inventario y almacenamiento.
- **Aprovechar mejor los recursos** disponibles en el área de planificación de inventario (Planning y/o logística dependiendo del área dedicada).

2.2 Objetivos de Investigación

- **Objetivo General**

Desarrollar un modelo de predicción de ventas de repuestos automotrices que ayude y oriente a la gestión del inventario, de esta forma permita equilibrar la disponibilidad de productos y los costos asociados, con una visión tanto de la utilidad como del valor total agregado.

- **Objetivos Específicos**

1. **Analizar y caracterizar** los datos históricos de ventas de repuestos automotrices, destacando la naturaleza intermitente o errática de algunos ítems.
2. **Comparar diferentes enfoques predictivos** (modelos clásicos de series de tiempo vs. modelos de machine learning basados en árboles de decisión y así mismo el modelo Croston).

3. **Diseñar y aplicar** una métrica de utilidad que considere costos de inventario, quiebres de stock e ingresos por ventas.
4. **Evaluar y discutir** los resultados de cada modelo, identificando fortalezas y las limitaciones que se encontraron.
5. **Proponer mejoras y líneas de investigación futuras**, incluyendo consideraciones éticas.

3. Estado del Arte

La literatura reciente en **forecast de demanda de repuestos** destaca la complejidad de predecir ventas con comportamientos **intermitentes** o **erráticos**, así como en la necesidad de ir más allá de las métricas de regresión tradicionales para optimizar la gestión de inventario. Hay numerosas formas de abordar este problema. Entre las diferentes fuentes que se consultaron respecto al estado del arte que aborda este problema se encuentran las siguientes:

3.1 **Benchmarking Spare Part Demand Forecasting Methods** (De Haan, 2021)

- Caracteriza la demanda de repuestos según patrones (suavizada, grumosa, intermitente, errática).
- Compara diferentes métodos de pronóstico y destaca la importancia de métricas enfocadas en control de inventario.

3.2 **Improving Spare Parts Demand Forecasting at ASML with Installed Base Information** (Van den Oord, 2022)

- Introduce el enfoque de **Installed Base**, que considera la probabilidad de falla de un repuesto desde su instalación.
- Sugiere la relevancia de combinar información sobre la vida útil de los repuestos con datos de venta.

3.3 **Spare Parts Demand Forecasting Method of Modern Enterprises Based on Digital Twin Model** (Wang et al., 2022)

- Propone un “gemelo digital” que integra factores operacionales y la probabilidad de vida útil de los repuestos.
- Resalta la utilidad de modelos de pronóstico no solo matemáticos, sino basados en el comportamiento real de los productos.

3.4 **Intermittent Demand Forecasting for Spare Parts: A Critical Review** (Pinçe et al., 2021)

- Analiza métodos paramétricos y no paramétricos para **demandas intermitentes**.
- Discute la complejidad de predecir ventas con datos escasos y Peaks no regulares.

3.5 Spare Parts Demand Forecasting in Maintenance, Repair & Overhaul (Lucht et al., 2022)

- Explora enfoques de machine learning para la industria automotriz, enfatizando la precisión de los pronósticos y la integración con la planificación de materiales.

3.6 Classification for Forecasting and Stock Control: A Case Study (Boylan, 2008)

- Introduce umbrales heurísticos para la clasificación de la demanda (parámetros p y CV^2), valiosos para identificar qué modelos son más apropiados según el tipo de demanda.

La revisión de estos trabajos deja en evidencia que el **control de inventario** es un factor clave en la evaluación de los modelos de forecast y que las demandas no siempre se comportan de forma regular o estable. En particular, en muchos contextos suele ser bastante errática tal y como el problema que abordamos, en el que en concreto tenemos la mayoría de los períodos como venta nula y en muy pocos tenemos venta de cada repuesto. Además, se aprecia la necesidad de profundizar en la creación de **métricas económicas o de utilidad** (más allá de la precisión estadística del modelo) que cuantifiquen el impacto real sobre el negocio medido en efectos concretos o KPIs, basándonos en la resta de ingresos (ventas de repuestos) menos costos por quiebre de stock y/o inventario en exceso de repuestos.

4. Resumen Técnico del Proyecto (Ciclo de Vida de un Proyecto de Ciencias de Datos)

Para llevar a cabo la investigación, se utilizó la **metodología CRISP-DM**, en este caso adaptada a los objetivos de optimización de la gestión de inventario:

4.1 Business Understanding

- Identificar la problemática de **costos por sobrestock y quiebres de stock**.
- Establecer el objetivo de desarrollar un modelo de **forecast** con métricas tradicionales y de utilidad.

4.2 Data Understanding

- **Datos:** Ventas reales de repuestos de empresa desde 2018-08 hasta 2024-09, incluyendo fecha de venta, ID del repuesto, cantidad vendida y precio unitario. Estos datos fueron anonimizados para efectos de este trabajo.
- Limpieza de datos y **relleno de periodos sin ventas** con registros de venta = 0 para capturar la intermitencia.

4.3 Data Preparation

- Clasificación de ítems según tipo de demanda (suavizada, grumosa, intermitente, errática). [Boylan, 2008; De Haan, 2021].
- Creación de **features** de tiempo (día de la semana, mes, año, etc.) y variables que representen la ID del repuesto.

4.4 Modeling

- **Modelos de series de tiempo:** (p. ej. SARIMAX) para capturar estacionalidad y correlaciones temporales y Croston para demanda errática.
- **Modelos de machine learning** (RandomForestRegressor, XGBoostRegressor) para capturar relaciones más complejas e interacciones entre variables.

4.5 Evaluation

- **Métricas clásicas de error** (MAE). Esto es el promedio de la suma del error absoluto entre dato observado de ventas versus el pronóstico de ventas.
- **Métrica de utilidad:**

Supuesto utilizado: Costo de inventario fijo: A modo de simplificar la cuantificación del costo de inventario, consideramos este costo como un costo fijo y similar para todos los productos.

Componentes de la métrica de utilidad:

- **Ingreso por venta:** Por cada repuesto
Ingreso en periodo i = Suma (# venta repuesto j x precio repuesto j en periodo i)
- **Costo de inventario:** este costo está dado cuando la cantidad predicha de demanda de repuesto es mayor a la cantidad observada en el set de evaluación. Se consideró el **precio unitario promedio** de todos los productos multiplicado por 0.3 (30% del precio promedio, dado por la posible depreciación del producto).
 - **Precio promedio repuestos:** 2,3749
 - **Costo inventario:** 0.71248

Costo inventario en periodo i = Suma (# exceso repuesto j x costo inventario en periodo i)

- **Costo por quiebre de stock (stock out):** se incurre en este costo cuando la cantidad predicha de demanda de repuesto es menor que la cantidad observada de venta, por lo que se considera el Es el costo de oportunidad de no haber podido vender el delta de repuestos faltantes en un periodo dado.

Costo quiebre stock en periodo i = Suma (# repuesto j faltante x precio repuesto j en periodo i)

Utilidad:

Suma Ingresos periodo i - Suma (costo inventario periodo i - costo stock out periodo i)

4.6 Deployment

- Implementación en entornos de prueba (Ej. planillas interactivas o dashboards).
- Validación y medición continua para mejorar los modelos.

- Uso de períodos acotados para el pronóstico futuro de venta. Mientras más lejano es el pronóstico, menor precisión tiene el modelo para determinar ventas.
- La toma de decisión debe estar acompañada por criterio experto y basado en **criterios éticos**.
 - No se debería dejar la total responsabilidad de planificación de ventas futuras al modelo. Los datos van cambiando su distribución y comportamiento a lo largo del tiempo, por ende, modelos deben ser entrenados siempre con datos actualizados.
 - Se debe considerar el modelo como un apoyo a la toma de decisión, pero se debe acompañar de criterio experto y con enfoque práctico basado en experiencia del equipo de logística y ventas.
 - Las personas encargadas de logística y ventas no deberían ser perjudicados o reemplazados en el alcance de su rol o responsabilidad por los modelos. Estos no fueron pensados con este fin.

5. Resultados

Para hacer frente al problema se decidió utilizar un set de entrenamiento y un set de validación, en una proporción de 80:20. Para lo anterior, como comentamos anteriormente, se rellenaron los periodos diarios de venta nula para cada repuesto entre la primera y última venta del repuesto.

Dataset	Total datos
Entrenamiento	2.670.194
Validación	667.549

Por otro lado, se consideró entrenar todos los modelos en torno a periodos de venta diaria, agregando la venta por día y repuesto, para los casos en que ocurría más de una venta de repuesto por día. Asimismo, como se ha comentado antes, se consideró rellenar los períodos de venta nula entre la primera y última venta para cada repuesto.

Finalmente, a modo de medir la calidad de ajuste de cada modelo en torno a las métricas definidas, se midieron estas de forma mensual, agregando la venta de cada repuesto de forma mensual.

Primer enfoque: Predicción binarias de venta, luego regresión lineal para cantidad de ventas

El primer enfoque que decidimos utilizar fue estructurar un modelo de clasificación binaria, en el que para cada repuesto se predecía si iba a existir demanda por el repuesto (valor 1), o el caso contrario en que no hay demanda (valor 0) en cada uno de los períodos de posible venta del repuesto. Posterior a esta clasificación, se utiliza un modelo de regresión lineal para predecir, por

cada repuesto y cada periodo con resultado positivo (hay venta), la magnitud de ventas que se experimenta.

Este enfoque fue rápidamente **descartado**. Los resultado de la predicción binaria de si existirá venta eran similares a aplicar un modelo aleatorio o incluso de peor rendimiento, dado que para cada repuesto existen demasiados períodos nulos de venta entre la primera y última venta, lo que conduce a una situación de clases desbalanceadas, en las que el modelo tiende a clasificar como un caso de venta nulo en prácticamente todo el set de validación. Con esto teníamos métricas de *accuracy* altas pero métricas de *recall* y *precision* muy bajas (menores a 0.1), lo que es común en este tipo de escenarios. Considerando este problema y otros más, asociados a aplicar modelos de caja negra, estando en un contexto de empresa y decisiones en que requiere explicar las decisiones del modelo, es que **se decidió descartar de plano este primer enfoque**.

Segundo enfoque: Machine Learning sobre datasets con variables de tiempo.

Dado este enfoque se decidió crear **variables en base al periodo** que se realizó la venta, incluyendo el día, mes, año y día de la semana como variables separadas, así como **variables contextuales de si la venta fue un día de semana o fin de semana**, y una **variable estacional** asociada a la cantidad de días transcurridos desde la primera venta de dicho repuesto.

Nombre variable	Tipo variable
ítem	numérica
day	númerica
month	númerica
year	númerica
weekday	númerica
is_weekend	booleana
days_since_first_data	númerica
sales (TARGET)	númerica

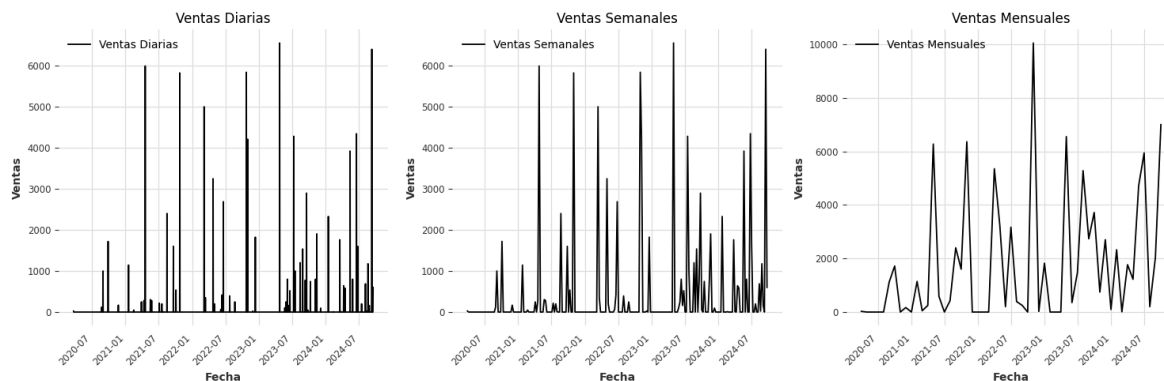
Esquema de dataset utilizado en modelos. Fuente: Elaboración propia.

Se entrenaron dos modelos basados en árboles de machine learning para la predicción de **cantidad de ventas futuras de repuestos**. El primer modelo de ensamble de árboles; *Random Forest*, el segundo modelo de *gradient boosting*; *XGBoost*. Más adelante mostraremos los resultados comparativos de las métricas que arrojaron estos modelos, en conjunto a los demás modelos utilizados para enfrentar el problema.

Tercer enfoque: uso de modelo de Croston para demanda intermitente

Se decidió utilizar el modelo de Croston para predecir la cantidad de ventas de repuestos, bajo dos datasets diferentes: sin estacionalidad y con estacionalidad.

El modelo de Croston es un modelo que utiliza como fundamento el alisado exponencial utilizado en predicción de series de tiempo frente a **escenarios de demandas intermitentes**, por ejemplo, como se puede apreciar en la siguiente figura, un producto que presenta muchos periodos intermedios con valor nulo y ventas intermitentes, como es el caso de las ventas de repuestos que tratamos.



Comportamiento de ventas según diferente temporalidad para el repuesto de ID 6815. Fuente: Elaboración propia

La funcionalidad del modelo consiste en dos componentes:

- Promedio de las ventas no nulas: Donde se calcula el promedio de las demandas que realmente ocurren, ignorando los periodos en los que no hay demanda.
- Promedio de los intervalos entre las demandas no nulas: Estima el tiempo promedio entre eventos de demanda.

Este modelo estima la demanda futura como el producto de ambos componentes, siendo más adecuado que otros métodos para productos con una demanda errática.

Se decidió entrenar un **modelo de Croston solo en base al periodo** de la venta del ítem y también se entrenó un **modelo de Croston que incorpora un factor estacional** en el que se calcula un ratio entre las ventas de dicho mes sobre las ventas totales, para cada mes del año.

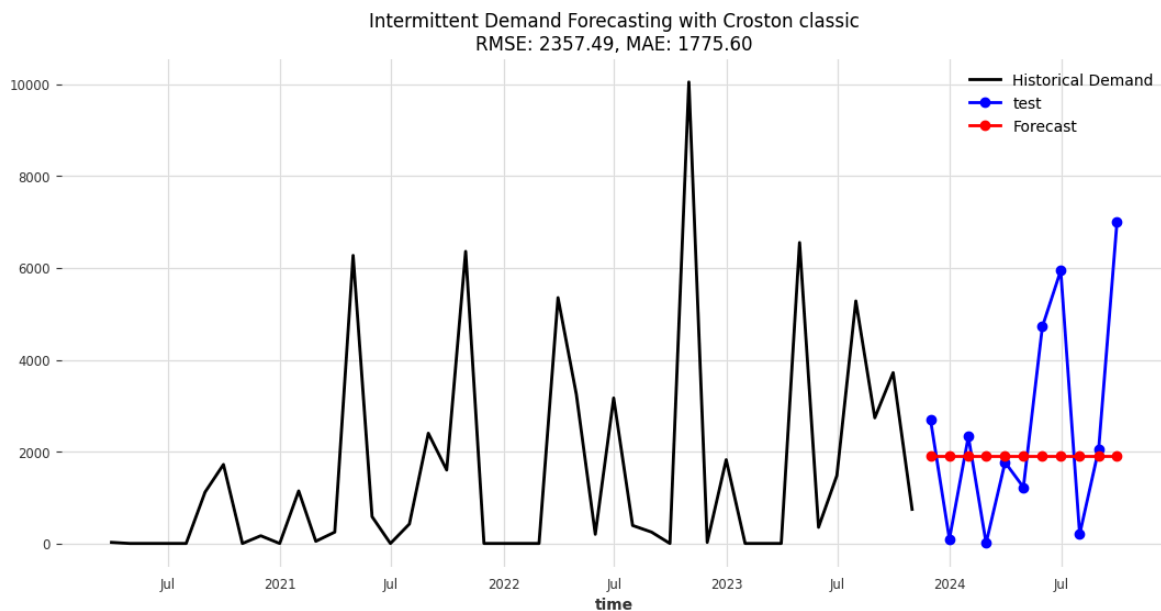
Tabla 1. Resultados Comparativos

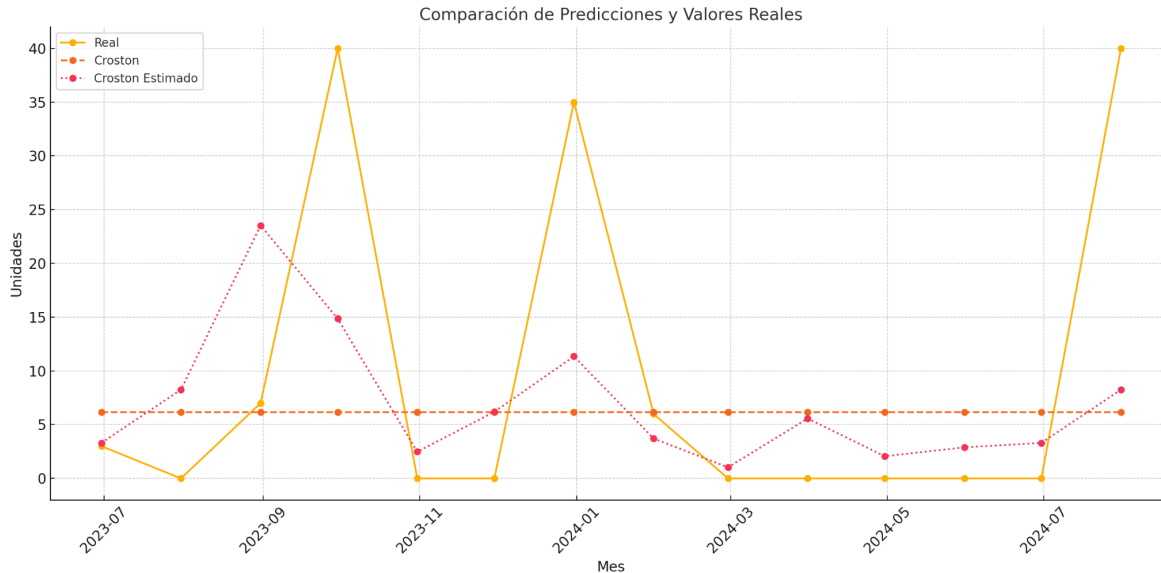
Modelo	MAE	Métrica utilidad (CLP)
RandomForestRegressor	7.75	-491,581.23
XGBoostRegressor	8.16	-485,878.72
Croston	5.1645	-1,235,960.13

Croston con estacionalidad	2.6029	-426,611.02
----------------------------	--------	-------------

- **MAE:** El *Random Forest* presenta un MAE ligeramente menor al modelo *XGBoost*, lo que indica un error promedio algo más reducido pero que no entrega resultados muy concluyentes. Por otro lado, se observa que los modelos de Croston tienen un MAE notoriamente mejor, sobre todo en el caso de Croston con estacionalidad.
- **Métrica de utilidad:** Los modelos presentan un valor negativo (pérdidas estimadas) en la simulación, reflejando costos acumulados más altos que las ganancias. Entre los dos modelos de Machine Learning, *XGBoost* muestra una menor pérdida total (-495.630,66 vs. -505.168,64) lo cual refleja que son bastante similares, mientras que en el caso de Croston se observa que el modelo con estacionalidad tiene un escenario de utilidad más favorable que el resto.

Para tener un entendimiento más gráfico del funcionamiento del modelo, a continuación se presentan 2 figuras en las que se ve claramente como el modelo de croston predice los valores futuros con una estimación en base a la cantidad de períodos con valor 0 y los valores no nulos.





Se puede observar que el modelo de Croston Estacionario refleja de una forma mucho mejor el componente de periodicidad del comportamiento de la demanda periódica. Por otro lado, el modelo de Croston estándar tiene dificultades para capturar patrones estacionales en los datos. Por esta razón, se extendió su uso mediante ajustes estacionales basados en promedios mensuales.

6. Discusión de Resultados

Los resultados muestran que, si bien los modelos de **árboles de decisión** (RandomForest y XGBoost) ofrecen mejoras en algunos escenarios (por ejemplo, en datos con suficiente densidad), ambos encuentran limitaciones en la **demanda intermitente** y peaks aislados. Por otro lado, los modelos de Croston ofrecen una métrica mucho mejor de calidad de ajuste en torno al error, si bien siguen teniendo problemas en la métrica de utilidad, es particularmente bastante bueno en abordar el problema de demanda intermitente.

Entre los factores más importantes a destacar:

6.1 Fortalezas del modelo de Croston

- El modelo de Croston, al basarse en un modelo de suavizamiento exponencial, es bastante útil para predecir demandas intermitentes, generando de esta forma mayor consistencia a nivel mensual. Esto resulta bastante positivo, ya que el modelo no solo proyecta la demanda de los ítems, sino que captura la demanda acumulativa mensual de estos, cubriendo de esta forma los peaks que se puedan presentar a futuro.

6.2 Modelos tienen utilidad muy negativa

- Esta situación es esperable, sobretudo en los modelos de Croston, ya que realizan un suavizamiento de la demanda, teniendo en el pronóstico muchos periodos más de demanda de repuestos no nula. Lo anterior genera costos sobretudo de inventario

(cantidad predicha > cantidad vendida observada), los que genera entonces utilidad negativa por excesivo inventario.

6.3 Falta de correlación clara entre precio y demanda

- El mercado de repuestos no siempre responde de manera directa a variaciones de precio. Factores como la urgencia de reparación y la obsolescencia del modelo de vehículo pueden tener mayor influencia.

6.4 Métrica de utilidad

- Si bien las pérdidas en la simulación fueron elevadas, la introducción de esta métrica aporta **una perspectiva realista** de los costos relacionados con la gestión de inventario que es lo que se quiere maximizar.
- Se identificó que ciertas ventas muy altas (con predicciones muy altas) conllevan costos de inventario significativos, mientras que subestimaciones graves generan un alto costo por quiebre de stock.

6.5 Comportamiento no esperado en datasets

- El id de cada repuesto, al tener un valor numérico, puede derivar en que el modelo asocie magnitudes mayores o menores dependiendo del repuesto en cuestión, lo que en consecuencia genera el aprendizaje de patrones que no tienen que ver con la realidad de los datos.

6.6 Implicaciones para la toma de decisiones

- Las empresas deben calibrar sus modelos no solo para minimizar el error promedio (MAE), sino para maximizar la utilidad o los KPIs que se definan, lo que puede implicar aceptar un error algo mayor si con ello se reducen costos de sobrestock y quiebres que es lo que se quiere evitar.

6.7 Interpretación de resultados

- Para los resultados obtenidos con Croston, es importante entender que debido a la intermitencia que presentan los datos y que el modelo se basa en exponential smoothing o alisamiento exponencial, este entrega un resultado mensual constante y estima un stock acumulativo mensual para abarcar los peaks de demanda que se presentan en los periodos siguientes.

En definitiva, los resultados concuerdan con la **literatura** que destaca la complejidad de la demanda intermitente (Pinçe et al., 2021) y la relevancia de introducir elementos de negocio en las métricas de evaluación (Lucht et al., 2022). Sin embargo, se abren algunas oportunidades para ajustar los modelos a subgrupos de repuestos o explorar metodologías híbridas o como en este caso Croston que tuvo los mejores resultados especialmente con ajustes estacionales, mostrando un rendimiento robusto, adaptándose de mejor forma a demandas intermitentes. Es por ello que para este tipo de demanda, el modelo de Croston con datos estacionales se presenta como una de las mejores alternativas disponibles en cuanto a predicción de demanda intermitente mensual.

7. Conclusiones

- **Conclusión:** Un **forecast de ventas de repuestos** que considere tanto métricas de error como una **métrica de utilidad** puede proveer un equilibrio más adecuado entre disponibilidad y costos de inventario.
- **Alineación con Objetivos:**
 - 7.1 Datos Históricos:** Logramos analizar y clasificar la demanda según distintos grupos (suavizada, grumosa, intermitente, errática).
 - 7.2 Comparación de Modelos:** Los modelos de machine learning (RandomForest, XGBoost) tuvieron un rendimiento aceptable en términos de MAE, pero la métrica de utilidad evidenció limitaciones y la necesidad de un ajuste más fino por ello finalmente se decantó por Croston, el cuál entregó mejores ajustes de ventas.
 - 7.3 Métrica de Utilidad:** La introducción de esta métrica permite evaluar el impacto económico de los errores de predicción, otorgando una alternativa para la elección de modelos que sea más ad hoc a la realidad de la empresa.
 - 7.4 Discusión de Resultados:** Identificamos la complejidad de predecir peaks en demanda intermitente y la falta de correlación directa con el factor precio. Se debe considerar incluir el comportamiento de la demanda de repuestos como variable predictora en el pronóstico de ventas.
 - 7.5 Aporte al estado del arte:** El estudio refuerza la importancia de metodologías híbridas y de la métrica de utilidad para un mejor **control de inventario**. En cuanto a la métrica de utilidad, se requiere afinar los factores en cuánto a los costos incurridos, acorde a la empresa que vaya a utilizar el modelo seleccionado, para obtener resultados económicos más realistas.

8. Análisis de Alcances Éticos del Proyecto

La ausencia de aspectos éticos puede afectar la reputación no solo de la empresa, sino también la de todos sus miembros, es por esto que es indispensable tener en consideración todos los alcances y las implicancias éticas al momento de crear un modelo. Por otro lado, es fundamental definir sólidamente estos aspectos para que en caso de que se busque replicar o mejorar este trabajo, se tengan en consideración todas las implicancias éticas que pueda suponer esto. En el aspecto ético que conlleva generar modelos para predecir el stock, se consideraron los siguientes puntos:

- **Automatización y Desempleo:** Al crear un modelo predictivo el cual tenga la capacidad de abarcar todas las estimaciones de stock necesarios para los periodos de la compañía, puede reducir la necesidad de trabajadores en áreas como la planificación de inventario y la logística. Esto puede incurrir en desplazamientos de empleados o despidos, lo cual tiene implicancias éticas en ámbitos como el desempleo y la justicia social. Es por esto que se debe crear un modelo moldeable, considerando que los modelos no son perfectos, para

que los involucrados en el manejo del stock logren potenciar su trabajo a través de esta herramienta y se obtenga el mejor provecho.

- **Condiciones de Proveedores:** Debido al surgimiento de modelos predictivos, se pueden generar mayores tensiones con proveedores debido a tener que crear cronogramas más estrictos y volúmenes más grandes de pedidos, así pudiendo afectar las condiciones laborales de ambas partes, que como lo haría un planificador que tendría más holgura y tendría en consideración otros ámbitos. Teniendo estos puntos en consideración, se debe planificar la implementación de este modelo, con marchas blancas para tener periodos de adaptación al cambio e informando a las partes involucradas los cambios que se harán, proponiendo nuevos tiempos que se irán adaptando a medida que el modelo sea utilizado.
- **Sesgo en los datos:** En algunos casos, las predicciones pueden estar diseñadas para maximizar las ganancias a corto plazo sin tener en cuenta el impacto a largo plazo en la comunidad o el medio ambiente. Las empresas deben equilibrar los objetivos comerciales con su responsabilidad sobre la empresa, por ejemplo, un gerente busca aplicar modelos que sean buenos a corto plazo para así generar buenos números y estadísticas, por otro lado un dueño o los inversionistas buscarán implementar los mejores modelos a largo plazo que sean realmente beneficiosos para la empresa.
- **Desechos de Partes No Utilizadas:** Las piezas automotrices tienen ciclos de vida más largos, por lo que, si no se venden, pueden convertirse en residuos difíciles de manejar. Un exceso de inventario debido a predicciones poco precisas puede llevar a generación de chatarra automotriz, lo que generaría problemas medioambientales. Con esto vendría un ajuste al modelo de tal manera que priorice la disminución de residuos o que al menos equilibre ganancia con sustentabilidad.

En resumen, la toma de decisiones basada en datos no solo debe enfocarse en la **eficiencia de la utilidad de la operación**, sino también en apoyar la **responsabilidad social** y la **sostenibilidad** de la empresa en el tiempo, además de tener la capacidad de ser replicable en el futuro para generar un real aporte a la ciencia de datos.

9. Futuros Desarrollos

A partir de las conclusiones y resultados de este proyecto, se proponen las siguientes líneas de investigación y desarrollo:

9.1 Modelos Especializados para Demanda Intermitente

- Incluir modelos mixtos que incorporen a **Croston** o **SBA (Syntetos-Boylan Approximation)** que se especializan bastante bien por lo que pudimos analizar en este tipo de problemas de demanda intermitente.

9.2 Integración de Fuentes Adicionales de Datos

- Información de marketing, campañas, vida útil real de los automóviles, o datos contextuales (tendencias macroeconómicas) para así tener mejores input para entregar el modelo y poder desagregar el efecto de ciertas variables.

9.3 Optimización a varios niveles

- Combinar la **precisión** (MAE, RMSE) con una **función de utilidad** (costos/beneficios) y con criterios de **sostenibilidad medioambiental**. Además, se pueden analizar diferentes segmentos de ítems con comportamiento similar y aplicar modelos o análisis diferenciados.

9.4 Sistemas de Retroalimentación Continua

- Implementar **Dashboards** interactivos que permitan a los equipos de inventario ajustar las predicciones según la disponibilidad real o eventos imprevistos (alertas tempranas de quiebres, etc.).

9.5 Aplicación de Juicio Experto

- La experiencia de los planificadores y técnicos de repuestos puede contribuir a afinar y mejorar el forecast, sobre todo en situaciones de cambio drástico en la demanda o consideraciones comerciales que puedan surgir.

Estas propuestas apuntan a una **adopción más amplia** de la analítica avanzada de datos en la industria automotriz, garantizando que los resultados se puedan replicar y asimismo que la generación de valor sea sostenible.

Referencias

1. Boylan, J. (2008). Classification for forecasting and stock control: A case study. *Journal of the Operational Research Society*, 59(4). <https://doi.org/10.1057/palgrave.jors.2602312>
2. De Haan, D. (2021). *Benchmarking Spare Part Demand Forecasting Methods*. Erasmus University Rotterdam.
3. Lucht, T., Aliksieiev, V., Kämpfer, T., & Nyhuis, P. (2022). Spare Parts Demand Forecasting in Maintenance, Repair & Overhaul. *Conference on Production Systems and Logistics (CPSL) 2022*, Vancouver, Canada. <https://doi.org/10.15488/12179>
4. Pinçe, Ç., Turrini, L., & Meissner, J. (2021). Intermittent Demand Forecasting for Spare Parts: A Critical Review. *Omega*, 105, 102513. <https://doi.org/10.1016/j.omega.2021.102513>
5. Van den Oord, W.D. (2022). *Improving Spare Parts Demand Forecasting at ASML with Installed Base Information*. Eindhoven University of Technology.
6. Wang, S., Wang, Y., & Wang, J. (2022). Spare Parts Demand Forecasting Method of Modern Enterprises Based on Digital Twin Model. *International Journal of Modeling, Simulation, and Scientific Computing*, 13(6). <https://doi.org/10.1142/S1793962322500453>.

Anexo

Código fuente:

https://drive.google.com/file/d/1xUjs3us4zy6HWCQ99ROrNyovkinSdbyy/view?usp=drive_link.