

Reporte de la Primera Etapa del Proyecto

Integrantes:

- Andrés Cano
 - Andrés Meneses
 - Jhonattan Reales
-

Proyecto: Pronóstico de la Demanda para Productos de Vida Útil Corta, en la categoría de líquidos, en una Empresa B2B de Derivados Lácteos

1. Análisis del Problema

1.1. Contexto La empresa objeto de estudio es una compañía B2B que comercializa productos lácteos, entre ellos una categoría crítica de productos líquidos con vida útil corta. Debido a la naturaleza perecedera de estos productos, la sincronización entre la producción, el despacho y la demanda real es esencial para evitar pérdidas por vencimiento o por falta de disponibilidad en los puntos de distribución.

1.2. Problema principal Actualmente, la empresa enfrenta dificultades para predecir la demanda real de los productos líquidos, lo cual genera:

- Exceso de inventario y deterioro de producto.
- Faltantes en puntos de venta que derivan en venta perdida.
- Ineficiencia en la logística de abastecimiento y distribución.
-

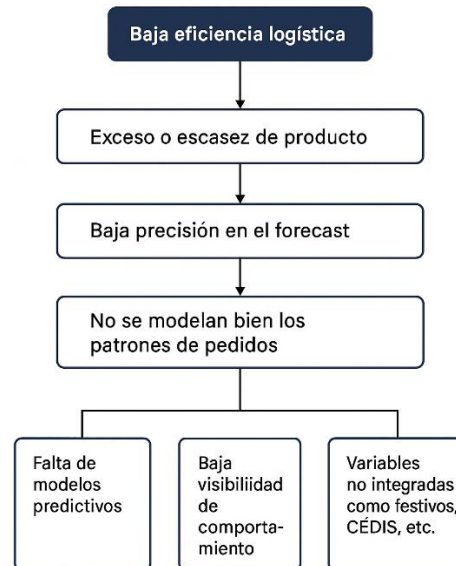
1.3. Objetivo General Desarrollar un modelo predictivo de demanda que permita superar el 65% de asertividad semanal en la categoría de productos líquidos, utilizando técnicas de Machine Learning y series de tiempo en un horizonte de 6 meses.

1.4. Objetivos Específicos

- Analizar patrones históricos de pedidos y envíos.
- Incorporar variables externas que impactan la demanda (festivos, quincenas, temporada escolar, etc.).
- Evaluar modelos de predicción tradicionales y de aprendizaje automático.
- Automatizar la generación de pronósticos semanales.

1.5. Pregunta SMART ¿Es posible superar el 65% de asertividad semanal del forecast en la categoría de líquidos en 6 meses aplicando machine learning para modelar patrones de compra del cliente?

1.6. Árbol de Problemas



2. Estado del Arte

2.1. Revisión bibliográfica

El pronóstico de demanda para productos perecederos o de corta vida útil, especialmente en el sector de alimentos y bebidas, representa un reto logístico y analítico importante. Esto puede ser resultado de la alta variación, vida útil limitada y patrones de consumo irregulares a los que se ven expuestos esta clase de productos.

Tradicionalmente, se han aplicado modelos estadísticos como Medias móviles, ARIMA y SARIMA para realizar estimaciones sobre datos históricos. Estos modelos han demostrado resultados aceptables en entornos donde la estacionalidad y la estabilidad son muy marcadas. Sin embargo, su rendimiento y precisión disminuyen cuando la demanda se ve afectada por múltiples variables externas o cuando los comportamientos históricos no se reproducen de manera constante.

Dado lo anterior, se ha impulsado el uso de modelos más adaptables y robustos. Un ejemplo de esto lo llevaron a cabo Taylor y Letham (2018), con el desarrollo de Prophet, una herramienta

de aprendizaje de máquinas que permite integrar eventos especiales y múltiples estacionalidades, resultando muy útil para productos cuya demanda se ve fuertemente influenciada por factores externos, tales como, días feriados, aspectos climáticos o ciclos de mercado, restricciones económicas, medidas políticas, entre otros. Prophet ha ganado popularidad en empresas que necesitan generar pronósticos precisos con bajo costo computacional y una interfaz accesible.

De manera similar, el aprendizaje automático ha cobrado relevancia en el contexto del pronóstico de demanda y/o la estimación de series temporales. Algoritmos como Random Forest, Gradient Boosting y XGBoost han sido implementados con éxito en cadenas de suministro de diversas industrias. En comparación con los algoritmos tradicionales, estos modelos se caracterizan por no asumir una estructura previa en los datos y son capaces de comprender relaciones no lineales complejas. Kamble et al. (2019) determinaron que el modelo XGBoost lograba una mejora notable en la precisión de los pronósticos, después de incorporar al modelo variables como el número de puntos de venta, días hábiles, feriados y tendencias históricas de pedidos.

Casos reales como los de Danone, Nestlé y Unilever demuestran que integrar datos históricos, factores contextuales y modelos automatizados ha permitido mejorar y optimizar la distribución de productos lácteos y disminuir notablemente el desperdicio. Según Kaur y Singh (2022), la adopción de modelos de machine learning en esta industria ha permitido alinear de mejor manera la producción con la demanda estimada, dando como resultado beneficios tanto económicos como ambientales.

Un ejemplo en América Latina es el estudio de Jaimes Campos y López Zúñiga (2021), quienes diseñaron un modelo de pronóstico para estimar la demanda semanal de productos alimenticios de alto consumo en un entorno real. Emplearon algoritmos de machine learning como XGBoost y Prophet, logrando mejorar de forma notable la precisión en comparación a los algoritmos tradicionales utilizados por la empresa del estudio. En este caso, los autores incorporaron variables como promociones, estacionalidades, feriados y patrones históricos de demanda.

Otro enfoque muy similar en el campo del aprendizaje de máquinas, es el uso de redes neuronales recurrentes, en particular LSTM (Long Short-Term Memory). Estas redes están diseñadas para captar dependencias a largo plazo en series temporales, lo que resulta útil cuando existen rezagos o efectos acumulativos en el comportamiento del consumidor. Estudios como el de Chien et al. (2021) aplicaron LSTM a productos alimenticios perecederos y obtuvieron mejoras en la precisión del pronóstico frente a otros modelos.

Además de los modelos en sí, la literatura enfatiza la importancia del preprocesamiento y la ingeniería de variables. La calidad del forecast no depende solo del algoritmo, sino de la riqueza y relevancia de las variables incluidas. Factores como si el día es feriado, si es fin de semana, el promedio de ventas históricas por punto de distribución, la cercanía al cierre de quincena, y la actividad escolar, son variables que pueden capturar patrones de comportamiento clave en la demanda.

En conclusión, los estudios revisados evidencian como un enfoque híbrido, combinando modelos estadísticos clásicos con técnicas modernas de aprendizaje automático, ofrece resultados significativos en la predicción de productos de vida útil corta. La clave del éxito reside en una integración efectiva entre los datos, las variables externas y la tecnología empleada.

BIBLIOGRAFIA

- Kamble, S.S., Gunasekaran, A., Gawankar, S.A. (2019). *Big data analytics for sustainable agricultural supply chains: A review and research framework*. Computers & Operations Research.
- Taylor, S.J., Letham, B. (2018). *Forecasting at scale*. The American Statistician.
- Chien, C.F., et al. (2021). *A hybrid LSTM-based demand forecasting model for perishable food products*. Journal of Food Engineering.
- Chopra, S., Meindl, P. (2020). *Supply Chain Management: Strategy, Planning, and Operation*. Pearson.
- Kaur, J., Singh, R. (2022). *Machine Learning Applications in Perishable Goods Supply Chain: A Case Study in Dairy Industry*. International Journal of Logistics Research and Applications.
- Jaimes Campos, D. L., & López Zúñiga, E. (2021). Modelo de forecast para predecir la demanda semanal de alimentos y bebidas de consumo masivo [Tesis de maestría, Universidad de los Andes]. Repositorio Institucional Séneca. <https://repositorio.uniandes.edu.co/handle/1992/64709>

3. Planeación del Proyecto

3.1. Lista de tareas

Recolección y limpieza de datos.	07/05/2025
Análisis exploratorio de datos (EDA).	14/05/2025
Ingeniería de variables (feature engineering).	21/05/2025
Entrenamiento de modelos de machine learning y series de tiempo.	29/05/2025
Evaluación comparativa de resultados.	13/06/2025
Generación automatizada de pronósticos.	18/06/2025
Elaboración del reporte final y video de presentación.	24/06/2025