Predicción de Tiempos de Entrega en la Cadena de Suministro: Una Solución de Machine Learning para la Gestión Eficiente del Inventario

Seminario de Analítica y Ciencia de Datos

# Resumen Descriptivo del Proyecto

Este proyecto tiene como objetivo desarrollar un modelo predictivo que, a partir de datos históricos y técnicas de machine learning, permita estimar con precisión los tiempos de entrega de los proveedores en la cadena de suministro. Con esta solución, buscamos enfrentar el desafío de la falta de materias primas importadas, que provoca interrupciones en los procesos. El modelo tendrá impacto sobre la gestión del inventario y reducirá los tiempos de inactividad, mejorando así la eficiencia operativa.

Marco Teórico:

## Revisión de Literatura.

La predicción de tiempos de entrega en la cadena de suministro ha sido objeto de extensa investigación en el campo de la gestión de operaciones y la analítica de datos. Dentro del marco de este proyecto, nos apoyaremos en referencias clave que aporten ideas significativas al desarrollo del proyecto.

El trabajo Banerjee et al. (2015) aborda la predicción de tiempos de entrega de proveedores en la cadena de suministro de piezas de motores de aeronaves, un problema similar al planteado. Para abordar este problema de tiempos de entrega de materia prima el documento desarrollo un método que combina modelos supervisados tal como regresión lineal y modelo ARIMA con una distribución gamma multivariante, estos modelos aprenden de datos históricos de las órdenes de compra culminadas, este enfoque permite modelar tanto factores continuos como categóricos y demuestra una alta efectividad minimizando los errores de predicción al prever los retrasos.

Alguno de los métodos utilizados en el documento es la regresión paso a paso el cual se emplea para identificar los factores predictivos más relevantes, por otro lado la distribución gamma multivariada útil para modelar tiempos de espera mejorando la precisión y se evaluó la eficacia de modelo mediante métricas de error de predicción y análisis de sensibilidad y especificidad. Aplicar estos métodos en el proyecto planteado puede ayudar a desarrollar un modelo predictivo robusto que mejore los tiempos de entrega de los proveedores reduciendo la incertidumbre en la disponibilidad del material.

Para el documento Steinberg et al. (2023) proponen un modelo de machine learning para predecir retrasos en la de proveedores en la cadena de suministro, especialmente en entornos de bajo volumen y alta variabilidad de la demanda. Este tipo de producción, caracterizado por una alta variabilidad y complejidad en componentes, presenta desafíos, como problemas de dimensionalidad el cual el proyecto aborda mediante métodos de aprendizaje supervisado y de regresión.

Alguno de los métodos utilizados es la regresión supervisada el cual estima el retraso en días calendario, también el estudio evalúa algoritmos como arboles de decisión, random forest, y redes neuronales, comparando su precisión en distintos puntos del proceso de compra, así mismo utiliza un control de la dimensionalidad generando codificación binaria para reducir la dimensionalidad sin excluir datos. Como resultado Este proyecto contribuye a una alta aplicabilidad en la cadena de suministro, permitiendo manejar datos de alta dimensionalidad sin comprometer a precisión así mismo se encontró que los algoritmos de boosting y random forest era los más efectivos para el caso, alcanzando altos niveles de precisión, también se aplicó el marco crisp que proporciona una estructura estandarizada para proyecto de análisis de datos.

El estudio de Zhou et al. (2023) aborda la estimación de tiempo de entrega mediante un modelo novedoso denominado inductive graph transformer, diseñado para plataformas de comercio electrónico donde los tiempos de entrega afectan directamente al cliente y los tiempos de compra, este modelo se centra capturar interacciones complejas entre las características de los pedidos tales como la dirección de origen, destino y la fecha de pago. Con el fin de hacer estimaciones precisas en los tiempos de entrega.

Utiliza un método de red convolucional sobre grafos, el cual permite que el modelo capture relaciones estructurales en los datos, como las conexiones entre minorista y la ubicación. También un transformador que actúa sobre la representación de los pedidos con el fin de capturar interacciones de alto orden entre sus elementos, permitiendo que el modelo tome decisiones más precisas. Además, utiliza el transformador como función de regresión para predecir los tiempos de entrega de manera inductiva.

La capacidad de la red convolucional para realizar predicciones complejas ayuda a predecir tiempos de entrega para relaciones previas como ubicaciones nuevas. Esto es valioso en procesos con cadenas de suministros o ubicaciones cambiantes ya que asegura la adaptabilidad del modelo a operaciones nuevasd.

El proyecto de Wolter y Hanne (2024) examina la predicción del tiempo de servicio en entregas a domicilio, para productos que requieren de montaje e instalación tipo muebles, utilizando algoritmos de machine learning. El objetivo es reducir la dependencia del conocimiento empírico del personal logístico específicamente planeando los tiempos de entrega y de instalación. El estudio utilizo redes neuronales artificiales y modelos de regresión, comparando su rendimiento en la estimación real de los tiempos de servicios frente a modelos como la regresión lineal multivariable y máquinas de soporte vectorial.

Como anterior mente se menciono se utilizaron algunas metodologías de algoritmos supervisados como maquinas de soporte vectorial que se usaron para comparar los resultados de clasificación y categorización de tiempos de servicio, por otro lado la regresión lineal múltiple que permitió analizar el rendimiento del modelo, sin embargo este modelo presento limitaciones en la precisión de los datos altamente variables, así mismo se utilizo tecinas de procesamiento de datos para limpiar y estructurar la información, lo que incluyo normalización, manejo de valores atípicos y codificación one hot para variables categóricas.

el modelo de ANN permite automatizar la predicción de tiempos de entrega y montaje, disminuyendo la dependencia del conocimiento personal. También la comparación de métodos demuestra que los modelos ANN y SVM pueden adaptarse a diferentes escenarios siendo flexibles en los procesos de analítica.

Maiti et al. (2014) propusieron un modelo basado en datos históricos para la predicción en tiempo real de la llegada de vehículos. Si bien este trabajo se centra en la predicción de tiempos de llegada de autobuses, su enfoque de utilizar solo un conjunto limitado de variables clave (trayectoria del vehículo y marcas de tiempo) para realizar predicciones precisas puede inspirar una arquitectura simplificada para el modelo de predicción de tiempos de entrega en la cadena de suministro. Este enfoque es particularmente útil cuando se busca un balance entre precisión y simplicidad operativa​ [5].

## Modelos y Métodos ya existentes.

[Céntrese en los métodos, modelos de ML o enfoques ya aplicados en proyectos similares al que están desarrollando. Se espera que ud analice las metodologías usadas previamente y justifiquen por qué algunas podrían ser útiles o descartables para su solución propuesta. También se espera que ud analice la originalidad de la que hablamos en clase y le invito a que incluya al CESET dentro de su base de datos de análisis]

## Diferenciación del Proyecto

[Elabore una discurso basado en los capítulos anteriores que muestre por qué su proyecto es innovador, ya sea por la vía de plantear algo totalmente nuevo o por la vía de mejorar significativamente algo existente. Sea recursivo, use su propio trabajo y el de los demás]

## Referencias.

[1] A. G. Banerjee, W. Yund, D. Yang, P. Koudal, J. Carbone, y J. Salvo, “A Hybrid Statistical Method for Accurate Prediction of Supplier Delivery Times of Aircraft Engine Parts”, en *Volume 1B: 35th Computers and Information in Engineering Conference*, Boston, Massachusetts, USA: American Society of Mechanical Engineers, ago. 2015, p. V01BT02A037. doi: 10.1115/DETC2015-47605.

[2] F. Steinberg, P. Burggräf, J. Wagner, B. Heinbach, T. Saßmannshausen, y A. Brintrup, “A novel machine learning model for predicting late supplier deliveries of low-volume-high-variety products with application in a German machinery industry”, *Supply Chain Anal.*, vol. 1, p. 100003, mar. 2023, doi: 10.1016/j.sca.2023.100003.

[3] X. Zhou, J. Wang, Y. Liu, X. Wu, Z. Shen, y C. Leung, “Inductive Graph Transformer for Delivery Time Estimation”, en *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*, feb. 2023, pp. 679–687. doi: 10.1145/3539597.3570409.

[4] J. Wolter y T. Hanne, “Prediction of service time for home delivery services using machine learning”, *Soft Comput.*, vol. 28, núm. 6, pp. 5045–5056, mar. 2024, doi: 10.1007/s00500-023-09220-7.

[5] S. Maiti, A. Pal, A. Pal, T. Chattopadhyay, y A. Mukherjee, “Historical data based real time prediction of vehicle arrival time”, en *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Qingdao, China: IEEE, oct. 2014, pp. 1837–1842. doi: 10.1109/ITSC.2014.6957960.

# Rúbricas

R01: Completitud y consistencia: el estudiante realiza una entrega del momento evaluativo con todos los elementos desarrollados, manteniendo una relación coherente entre los elementos solicitados. El estudiante usa una consistencia argumentativa que incluye el uso de referencias cuando hay mérito de autoría entre otros.

R02: Línea Base: el estudiante es capaz de establecer una clara línea de trabajos previos y teorías basadas alrededor de su proyecto de clase. Esto permite evidenciar que la solución propuesta en el trabajo precio es viable y tiene una base sólida teórica detrás de su posible solución.

R02: Innovación: dentro del ámbito de la analítica y la ciencia de datos el estudiante muestra, basado en la literatura existente, una clara diferenciación de su proyecto frente las soluciones tecnologico-metodológicas encontradas. Se tiene en cuenta la inclusión de la base de datos de CESET para la rúbrica.

R04: Referencias: hay referencias válidas según la descripción del capítulo y están adecuadamente vinculadas al estilo narrativo que permite una clara expresión de los resultados de la investigación de la línea base y el marco tecnológico-metodológico.