

Universidad de Lima  
Facultad de Ingeniería  
Carrera de Ingeniería Industrial



# **USE OF A MACHINE LEARNING MODEL FOR THE REDUCTION OF BACKORDERS IN THE CROSS DOCKING SALES PROCESS FOR THE HOMECENTER ORDER SERVICE**

Tesis para optar el Título Profesional de Ingeniero Industrial

**Jamil Venturo Panduro Lope**

**Código 20173692**

**Sebastian Emilio Pumayauri Hidalgo**

**Código 20173749**

**Asesor**

**Yvan Jesus Garcia Lopez**

Lima – Perú

Octubre de 2023



**Propuesta**  
**Carrera Ingeniería Industrial**

**USE OF A MACHINE LEARNING MODEL FOR THE REDUCTION OF BACKORDERS IN THE CROSS DOCKING SALES PROCESS FOR THE HOMECENTER ORDER SERVICE**

**Autor(es)**  
[20173692@aloe.ulima.edu.pe](mailto:20173692@aloe.ulima.edu.pe)  
[20173749@aloe.ulima.edu.pe](mailto:20173749@aloe.ulima.edu.pe)  
Universidad de Lima

**Resumen:** En el presente trabajo se precisa analizar el aumento de BackOrder en la atención de pedidos crossdocking en la atención de clientes Homecenter debido a la no definición de procesos de planeación de compras, derivando en costos logísticos cobros de fill rate y bajo nivel de servicio. De tal forma se pretende que las empresas que manejan alto volúmenes de inventario y pedidos constantes, deben contar con un plan de pronóstico para cubrir posibles quiebres de stock.

El principal propósito de la investigación es explicar una forma de prevenir roturas de stock usando un modelo de inteligencia artificial, basado en data histórica de ventas de una mediana empresa que maneja inventarios, así mismo se planteó el objetivo de determinar el modelo de machine learning para predecir y reducir los Backorders

Para el análisis de data se utilizó el Software Orange, donde se entrenó la data con diferentes modelos de inteligencia artificial como Árbol de decisiones, Máquina de soporte de vectores, bosque aleatorio, y redes neuronales. Donde se definió el modelo más preciso de acuerdo a indicadores numéricos como la matriz de confusión, el área bajo la curva (AUC) y el análisis de la curva ROC. Optando así por el modelo de redes neuronales, modelo que presentó datos más precisos.

Finalmente se presenta los resultados y se realiza la sugerencia a nivel de gerencia sobre la toma de decisiones en el proceso de abastecimiento. Para ello se considera pertinente ahondar en el tema de las variables influyen la acumulación de backorders

**Palabras Clave:** [Inteligencia artificial, pedidos pendientes, pronóstico de demanda e gestión de inventario

**Abstract:** In this work, it is necessary to analyze the increase of Back Order in the attention of crossdocking orders in the attention of Homecenter customers due to the lack of definition of purchase planning processes, resulting in logistics costs, fill rate charges and low service level. Thus, it is intended the companies that handle high volumes of inventory and constant orders should have a forecast plan to cover possible stock-outs.

The main purpose of the research is to explain a way to prevent stock-outs using an artificial intelligence model, based on historical sales data of a medium-sized company that manages inventories, as well as to determine the machine learning model to predict and reduce backorders.

For the data analysis, the Orange software was used, where the data was trained with different artificial intelligence models such as Decision Tree, Support Vector Machine, Random Forest, and neural networks. The most accurate model was defined according to numerical indicators such as the confusion matrix, the area under the curve (AUC) and the ROC curve analysis. Thus, we opted for the neural network model, which presented the most accurate data.

Finally, the results are presented and a suggestion is made at the management level regarding decision making in the supply process. For this purpose, it is considered pertinent to delve into the subject of the variables that influence the accumulation of backorders].

**Keywords:** [Machine learning, Backorders, Demand forecasting, Supply chain and inventory management].

**Línea de investigación IDIC – ULIMA**

**Área y Sub-áreas de Investigación:**  
[Supply Chain Management.]

**Objetivo (s) de Desarrollo Sostenible (ODS) relacionado (s) al tema de investigación.**

## PLANTEAMIENTO DEL PROBLEMA

Con la llegada del internet a finales del siglo XX, se masificó la información disponible para preparar modelos matemáticos de alta efectividad (Rudin, C., 2019). También la potencia del cálculo de las computadoras incrementó beneficiando a la automatización de procesos en diferentes empresas industriales y tecnológicas (Rudin, C., 2019).

En la industria tecnológica se aplica el concepto de “machine learning”, rama de la inteligencia artificial (IA), que permite el aprendizaje de las máquinas sin la necesidad de programarlas previamente para ello (Adadi, A. & Berrada, M., 2018). Actualmente, esta habilidad indispensable permite la sistematización para identificar patrones entre los datos y lograr predecirlos (Adadi, A. & Berrada, M., 2018)

Evidenciando la propuesta basada en tecnología por parte de esta industria, la necesidad de la automatización de procesos es latente, y por ello, se hace énfasis en abordar dicha temática mediante modelos que reduzcan los pedidos pendientes en procesos de ventas “cross docking”.

Por otro lado, también se menciona la mejora de atención de los pedidos “Homecenter”, cuya relevancia gira entorno al proceso logístico que agiliza la entrega de dichos pedidos. Es así como la problemática de este tema de investigación se centrará en la aplicación de “machine learning” para la gestión de inventarios y efectiva distribución de pedidos.

### OBJETIVOS

El objetivo de la investigación es determinar el modelo de machine learning para predecir y reducir los BackOrders,

### JUSTIFICACIÓN

El Backorder se define como aquel producto dentro de una orden de compra que no se puede atender por falta de stock, es decir que no se cuenta con la disponibilidad del producto por diferentes factores logísticos internos y externos, siendo así la generación de pedidos en espera por el cliente que en algunos casos se penaliza (Islam & Amin,2020). Por otro lado, el cross-docking es una técnica logística que se basa en reducir costos de inventario para aumentar la cantidad del flujo de mercancías o productos, de tal forma que se mejore la eficiencia de la cadena de suministro (Fonseca, et al, 2019). Las decisiones de programación son bastante relevantes para garantizar la aceleración en la rotación de inventarios y normalizar las entregas a tiempo (Fonseca, et al, 2019).

La definición del conjunto de instrucciones definidas por parte de los modelos de aprendizaje reforzado se diferencia en ser de clasificación y regresión (Higa, 2021). El primero tiene la desventaja de no ser adecuado para grandes conjuntos de datos debido a su alto tiempo de formación así mismo utiliza datos bidimensionales para su clasificación en el hiperplano, el segundo, sirve para predecir resultados en cantidad de acuerdo con el valor de las entradas o variables de análisis. (Higa, 2021)

### HIPÓTESIS

De acuerdo con la revisión sistemática de la literatura de investigación que ya fueron realizadas anteriormente. Se puede expresar la siguiente hipótesis: Es posible la reducción de Backorders en el proceso de atención de pedidos Homecenter aplicando un modelo de Machine Learning

## DISEÑO METODOLÓGICO

El presente trabajo de investigación tiene un enfoque cuantitativo y nivel explicativo-correlacional. Según la secuencia correlacional es longitudinal, ya que, se recopila data histórica de las variables logísticas (Pronósticos, Ventas, nivel de inventario, Backorder) entre los años 2018 hasta 2020 para ser analizadas, la secuencia longitudinal descompone data de un rango de tiempo que van a ser estudiados (Manterola et al., 2019).

Además, la metodología ofrece un producto final de explicación científica, debido al tipo de muestreo probabilístico, utilizando técnicas estadísticas para el análisis del comportamiento de las variables y sus dimensiones principales, además, Según Salgado M. (2019), el muestreo probabilístico analiza una parte de la población que cumple con la probabilidad de ser elegido al azar y se divide en tres tipos, estratificado, simple y sistemático.

La técnica de muestreo empleado es el estratificado, puesto que se dividirá el universo de productos en familias según el tipo de uso que se le da. El muestreo estratificado separa o subdivide a la población en grupos de características similares (Salgado M., 2019).

Seguidamente, el alcance de la investigación es de tipo explicativo porque se detalla las herramientas de inteligencia artificial que se utiliza y la estructura en que se compone los algoritmos.

La investigación explicativa se realiza para un problema que no se investigó bien antes, exige prioridades, genera definiciones operativas y proporciona un modelo mejor investigado. En realidad, es un tipo de diseño de investigación que se centra en explicar los aspectos de su estudio (“¿qué es la investigación explicativa?” 2020, parr. 1)

Para ello se recolecto data histórica de tiempo de tránsito, nivel de inventario actual, ventas y cálculo de pronóstico de ventas, a la cual se le denomina variables numéricas, seguidamente se identificó a las variables categóricas que tienen suma relevancia en la predicción del backorder.

## NOTAS (AGRADECIMIENTOS)

### REFERENCIAS

Adadi, A. & Berrada, M., (2018). Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). IEEE access, 6, 52138-52160. <https://doi.org/10.1109/ACCESS.2018.2870052>

Bronstein, M. M., Bruna, J., LeCun, Y., Szlam, A., & Vandergheynst, P. (2017). Geometric deep learning: going beyond euclidean data. IEEE Signal Processing Magazine, 34(4), 18-42. <https://doi.org/10.1109/MSP.2017.2693418>

Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O., & Walsh, A. (2018). Machine learning for molecular and materials science. Nature, 559(7715), 547-555. <https://doi.org/10.1038/s41586-018-0337-2>

Hajek, P., & Abedin, M. Z. (2020). A profit function-maximizing inventory backorder prediction system using big data analytics. IEEE Access, 8, 58982-58994. <https://doi.org/10.1109/ACCESS.2020.2983118>

Higa Shimabukuro, H. (2018, Enero). ANÁLISIS DE PRONÓSTICO DE DEMANDA PARA PODER GESTIONAR EL INVENTARIO APLICADO AL SECTOR SALUD

[https://repositorio.ulima.edu.pe/bitstream/handle/20.500.12724/14099/Higa\\_Shimabukuro](https://repositorio.ulima.edu.pe/bitstream/handle/20.500.12724/14099/Higa_Shimabukuro_Hugo.pdf?sequence=1&isAllowed=y)

\_Hugo.pdf?sequence=1&isAllowed=yIslam, S., & Amin, S. H. (2020). Prediction of probable backorder scenarios in the supply chain using Distributed Random Forest and Gradient Boosting Machine learning techniques. <https://link.springer.com/content/pdf/10.1186/s40537-020-00345-2.pdf>

Manterola, C., Quiroz, G., Salazar, P., & Garcia, N. (2019). recientemente utilizados en investigación clínica frecuentemente utilizados en investigación clínica.

Manterola, C., Quiroz, G., Salazar, P., & Garcia, N. (2019). recientemente utilizados en investigación clínica frecuentemente utilizados en investigación clínica.

<https://reader.elsevier.com/reader/sd/pii/S0716864019300057?token=022EF0A9DC5E61A2DEF905077DA6427A97A055AAE35FFDFDD5E7EEF2785F91796280E04A6729B32FEF05AC6401453D3D&originRegion=us-east-1&originCreation=20220626041125>

Marroquin J. M., Garcia Y. F, Taquia J. A. (2021). Machine Learning: Comparison of algorithms for determining water quality in the Rímac river. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(12), 552-572. <file:///F:/10%20CICLO/TESIS%201/3.%20VARIABLES/7406-Article%20Text-13459-1-10-20210526.pdf>

Namazkhan, M., Albers, C., & Steg, L. (2019, 29 de Octubre). A decision tree method for explaining household gas consumption: The role of building characteristics, socio-demographic variables, psychological factors and household behaviour. <https://doi.org/10.1016/j.rser.2019.109542>

Moshtagh, M. S., & Taleizadeh, A. A. (2017). Stochastic integrated manufacturing and remanufacturing model with shortage, rework and quality based return rate in a closed loop supply chain. *Journal of Cleaner Production*, 141, 1548-1573. <http://dx.doi.org/10.1016/j.jclepro.2016.09.173>

Chan, S. W., Tasmin, R., Aziati, A. N., Rasi, R. Z., Ismail, F. B., & Yaw, L. P. (2017, August). Factors influencing the effectiveness of inventory management in manufacturing SMEs. In *IOP Conference Series: Materials Science and Engineering* (Vol. 226, No. 1, p. 012024). IOP Publishing. <http://dx.doi.org/10.1088/1757-899X/226/1/012024>

Papernot, N., McDaniel, P., Goodfellow, I., Jha, S., Celik, Z. B., & Swami, A. (2017, April). Practical black-box attacks against machine learning. In *Proceedings of the 2017 ACM on Asia conference on computer and communications security* (pp. 506-519). <https://doi.org/10.1145/3052973.3053009>

Rudin, C., (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206-215. <https://doi.org/10.1038/s42256-019-0048-x>

Sarkar, B., Ahmed, W., Choi, S. B., & Tayyab, M. (2018). Sustainable inventory management for environmental impact through partial backordering and multi-trade-credit-period. *Sustainability*, 10(12), 4761. <https://doi.org/10.3390/su10124761>

Salgado M. (Septiembre de 2019). muestra probabilística y no probabilística. [http://ri.uaemex.mx/bitstream/handle/20.500.11799/108928/secme-10911\\_1.pdf?sequence=1](http://ri.uaemex.mx/bitstream/handle/20.500.11799/108928/secme-10911_1.pdf?sequence=1)

Tang, Z., & Ge, Y. (2021). CNN model optimization and intelligent balance model for material demand forecast. *International Journal of System Assurance Engineering and Management*, 1-9. <https://doi.org/10.1007/s13198-021-01157-0>

Zhang, M., Shi, L., Zhuo, X., & Liu, Y. (2021). A Bullwhip Effect Weakening Approach Based on VMD-SVM Algorithm under the Background of Intelligent Manufacturing. <https://doi.org/10.3390/pr9111957>

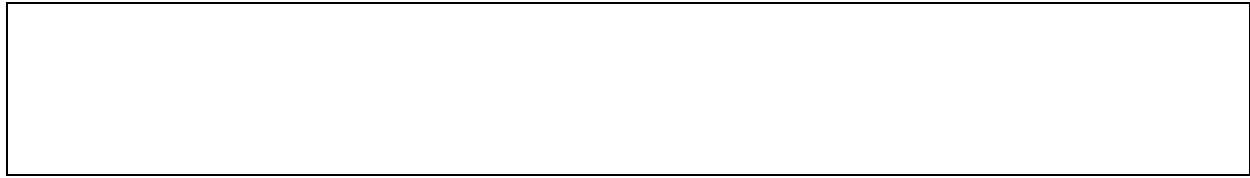
¿qué es la investigación explicativa? (6 de Julio de 2020). <https://tecnicasdeinvestigacion.com/investigacion-explicativa/>

Ojeda Riquezes, M. V. (2022). Sensibilidad y especificidad de una prueba comercial para la detección de *Candida albicans*.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., ... & Lee, S. I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature machine intelligence*, 2(1), 56-67. <https://doi.org/10.1038/s42256-019-0138-9>

Khan, M. Y., Qayoom, A., Nizami, M. S., Siddiqui, M. S., Wasi, S., & Raazi, S. M. K. U. R. (2021). Automated Prediction of Good Dictionary EXamples (GDEX): A Comprehensive Experiment with Distant Supervision, Machine Learning, and Word Embedding-Based Deep Learning Techniques. *Complexity*, 2021. <https://doi.org/10.1155/2021/2553199>

Fonseca, G. B., Nogueira, T. H., & Ravetti, M. G. (2019). A hybrid Lagrangian metaheuristic for the cross-docking flow shop scheduling problem. *European Journal of Operational Research*, 275(1), 139-154. <https://doi.org/10.1016/j.ejor.2018.11.033>

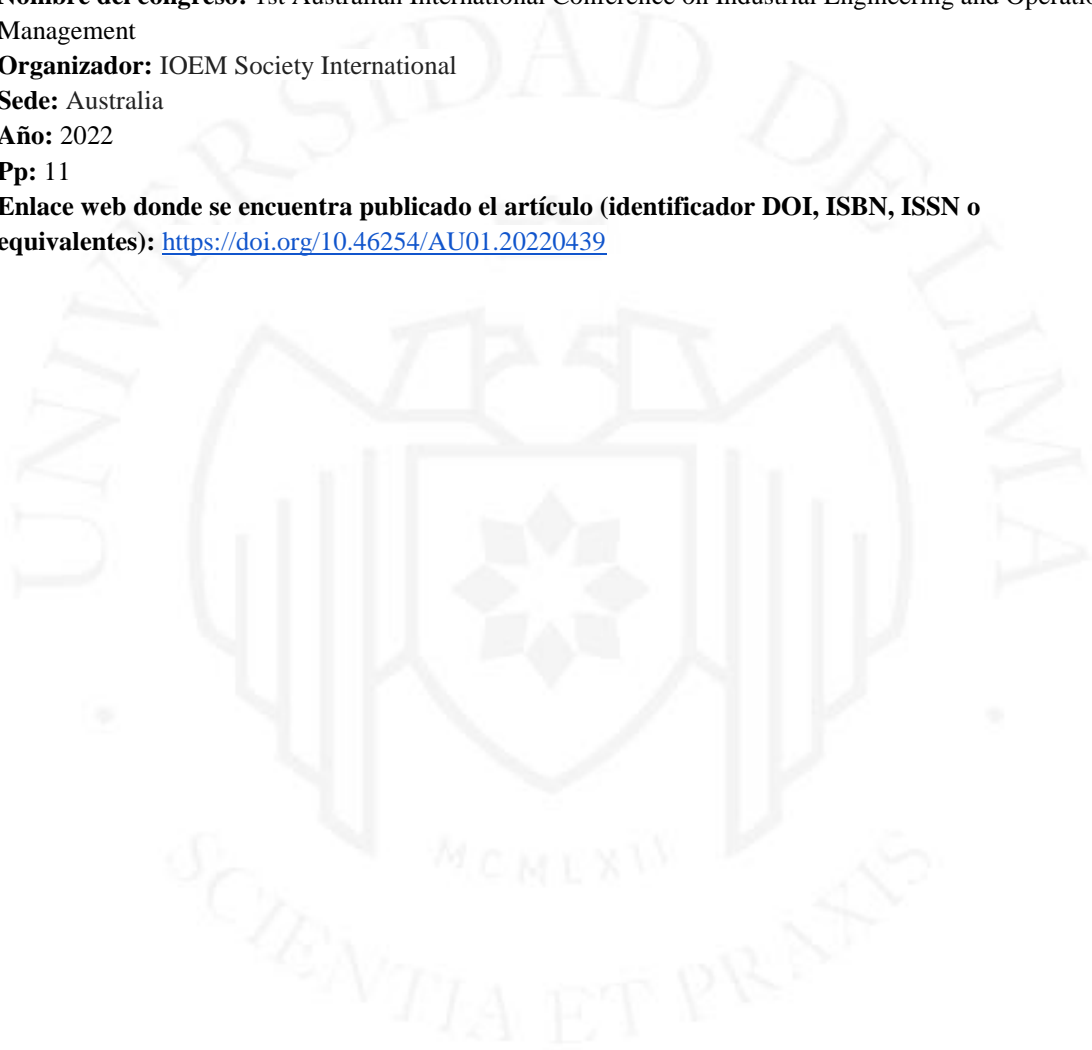


### **ANEXO. Datos del artículo publicado**

- **Nombre del artículo:** USE OF A MACHINE LEARNING MODEL FOR THE REDUCTION OF BACKORDERS IN THE CROSS DOCKING SALES PROCESS FOR THE HOMECENTER ORDER SERVICE
- **Autores:** Jamil Venturo Panduro Lope y Sebastian Emilio Pumayauri Hidalgo
- **Co autor(es):** Yvan Jesus Garcia Lopez.

#### **Presentación en congreso**

- **Nombre del congreso:** 1st Australian International Conference on Industrial Engineering and Operations Management
- **Organizador:** IOEM Society International
- **Sede:** Australia
- **Año:** 2022
- **Pp:** 11
- **Enlace web donde se encuentra publicado el artículo (identificador DOI, ISBN, ISSN o equivalentes):** <https://doi.org/10.46254/AU01.20220439>







## Paper IEOM

### INFORME DE ORIGINALIDAD

5%

INDICE DE SIMILITUD

3%

FUENTES DE INTERNET

4%

PUBLICACIONES

1%

TRABAJOS DEL ESTUDIANTE

### FUENTES PRIMARIAS

1

"Pervasive Computing and Social Networking", Springer Science and Business Media LLC, 2022

Publicación

2%

2

Submitted to Wilmington University

Trabajo del estudiante

1%

3

Smita Abhijit Ganjare, Sunil M. Satao, Vaibhav Narwane. "Systematic literature review of machine learning for manufacturing supply chain", The TQM Journal, 2023

Publicación

1%

4

[www.hindawi.com](http://www.hindawi.com)

Fuente de Internet

1%

5

Carlos Morales, Claudia Holtschlag, Aline D Masuda, Percy Marquina. "In which cultural contexts do individual values explain entrepreneurship? An integrative values framework using Schwartz's theories", International Small Business Journal: Researching Entrepreneurship, 2018

Publicación

<1%