Articles

Artificial intelligence demand forecasting for improved inventory and fleet management

Predicción de demanda mediante inteligencia artificial para mejorar la gestión de inventarios y flotas

Byron Oviedo-Bayas*
Elena López-Robayo*
Cristian G. Zambrano-Vega*

ABSTRACT

This study analyzes the impact of Artificial Intelligence (AI) on demand forecasting as a strategy to optimize inventory and logistics fleet management. Advanced models such as recurrent neural networks (RNN), Transformers, and Gradient Boosting were compared with traditional statistical methods such as ARIMA and Exponential Smoothing. The results showed that AI-based models reduced forecast error by up to 50%, decreased logistics costs by 31.5%, and reduced empty kilometers by 15%. In addition, the potential of AI to reduce CO_2 emissions was demonstrated, aligning with sustainability goals. However, significant challenges were identified, such as the need for high-quality data and specialized technical training. Future research lines focused on the scalability of AI for SMEs and its integration with blockchain are proposed.

Keywords: Time series, logistics sustainability, machine learning, route optimization, data quality.

RESUMEN

Este estudio analiza el impacto de la Inteligencia Artificial (IA) en la predicción de la demanda como

* Systems and Computer Engineer
Quevedo State Technical University
boviedo@uteq.edu.ec
* Systems and Information Technology Engineer
Quevedo State Technical University
elenamaribellopez@hotmail.com
* PhD in Systems and Information Technology
Quevedo State Technical University
czambrano@uteq.edu.ec

JOURNAL OF BUSINESS and entrepreneurial studies
ISSN: 2576-0971

Journal of Business and Entrepreneurship
July - September Vol. 9 - 3 - 2025
http://journalbusinesses.com/index.php/revista
e-ISSN: 2576-0971
journalbusinessentrepreneurial@gmail.com
Receipt: March 26, 2025
Approval: April 19, 2025
Page 1-9

estrategia para optimizar la gestión de inventarios y flotas logísticas. Se compararon modelos avanzados como redes neuronales recurrentes (RNN), Transformers y Gradient Boosting con métodos estadísticos tradicionales como ARIMA y Suavizado Exponencial. Los resultados mostraron que los modelos basados en IA redujeron el error de pronóstico hasta en un 50 %, disminuyeron los costos logísticos en un 31,5 % y redujeron los kilómetros en vacío en un 15 %. Además, se evidenció el potencial de la IA para reducir emisiones de CO_2, alineándose con objetivos de sostenibilidad. No obstante, se identificaron retos relevantes, como la necesidad de datos de alta calidad y la capacitación técnica especializada. Se propone explorar futuras líneas de investigación centradas en la escalabilidad de la IA para PYMEs y su integración con blockchain.
Palabras clave: Series temporales, sostenibilidad logística, aprendizaje automático, optimización de rutas, calidad de datos.

INTRODUCTION

Efficient inventory and transport fleet management is a key strategic element in optimizing complex and highly dynamic supply chains. In a global context marked by market volatility, the proliferation of distribution channels, increased competition, and the demand for operational sustainability, organizations face the constant challenge of aligning their logistics capacity with uncertain and changing demand (Jones, 2025).
Errors in demand forecasting lead to significant financial and operational impacts: excess inventory represents tied-up capital, increased storage costs, and the risk of obsolescence, while stockouts lead to lost sales, a poor customer experience, and contractual penalties. Several studies highlight that between 20% and 50% of logistics costs could be mitigated through more accurate demand forecasting models (Rocco et al., 2024).
Artificial Intelligence (AI), as an interdisciplinary field, has shown significant advances in the analysis of complex, nonlinear, and multivariate data. Machine learning algorithms allow hidden patterns, temporal interactions, and non-obvious correlations between variables to be captured, overcoming the limitations of conventional statistical models. According to Caso (2024), leading companies such as Amazon have achieved 95%

accuracy in their predictions by implementing deep neural networks, specifically RNN and LSTM, to estimate demand in their global distribution centers.

In the logistics environment, AI also allows the integration of exogenous data such as weather, promotions, macroeconomic events, urban mobility, and traffic patterns, which increases the adaptability of the forecasting system. For example, Li et al. (2024) applied deep learning-based models to predict urban deliveries in real time, reducing empty miles by 15% and improving fleet energy efficiency.

Models such as Transformers, initially developed for natural language processing tasks, have been successfully adapted to logistics time series due to their attention span and ability to model long-term dependencies (Wen et al., 2022). In 2023, Walmart used this architecture to anticipate weekly demand with error margins of less than 3%, representing a substantial improvement over classic statistical forecasting methods, which have errors of between 7% and 10%.

From an inventory management perspective, the use of AI has enabled organizations to adopt more agile and responsive schemes, such as just-in-time inventory, without sacrificing customer service levels. Amazon, for example, has integrated AI with computer vision and robotics systems in its logistics centers to forecast SKU-level demand, adjusting inventories in real time (Caso, 2024).

In terms of route planning, algorithms such as LSTM or CNN have been used to optimize dynamic fleet allocation, considering weather variables, delivery restrictions, and time windows. Ricci (2024) describes how FedEx incorporated reinforcement AI models to reconfigure its last-mile routes based on daily demand and weather conditions, improving vehicle occupancy by 18%.

In Latin America, the application of AI in logistics has been on the rise despite structural challenges. Mercado Libre has implemented gradient boosting algorithms in its real-time inventory management system, allowing it to adapt its warehouses to regional demand (Campos, 2017). In Chile, the startup Urzúa-Morales (2020) managed to reduce delivery times by 30% by implementing AI for urban route optimization.

Despite these achievements, significant barriers to the widespread adoption of AI in logistics remain: the lack of structured and labeled data, cultural resistance to automation, and the shortage of specialized talent are repeatedly cited as critical factors for failure. Rashid (2023) reports that 60% of AI projects in logistics fail to achieve their objectives due to deficiencies in data infrastructure and algorithmic governance.

On the other hand, the convergence of AI with emerging technologies such as blockchain, IoT, and digital twins promises new opportunities for traceability, security, and transparency in demand management. Charles et al. (2023) linked blockchain with predictive models to ensure the integrity of cold chains, improving data reliability in sectors such as pharmaceuticals and food.

In the academic sphere, multiple systematic reviews, such as those by Samun et al. (2025), have identified that hybrid models combining AI with mathematical optimization offer the greatest potential for solving complex logistics problems, particularly in the management of perishable products, where time and quality constraints are critical.

In summary, the application of AI in demand forecasting represents a radical transformation in modern logistics. From improving forecast accuracy to reducing CO_2 emissions, its benefits are manifold. However, successful adoption requires a robust data infrastructure, advanced analytical capabilities, and a long-term strategic vision.

This study contributes to the body of knowledge by empirically evaluating the effectiveness of AI models, specifically RNN, Transformers, and Gradient Boosting, compared to traditional techniques applied to demand forecasting and inventory and fleet optimization. Through a rigorous methodology and the use of real data, we seek to provide practical evidence on the impact of these technologies with a view to promoting their informed adoption in diverse operating environments.

MATERIALS AND METHODS

This study was developed using a quantitative and empirical approach, framed within applied research. This type of research was chosen because of its ability to generate useful knowledge for solving real problems, specifically in the context of demand forecasting and logistics optimization through the use of Artificial Intelligence (AI) algorithms. The objective was to evaluate, based on real data, the comparative effectiveness of traditional models and machine learning models in demand forecasting, as well as their impact on inventory and fleet efficiency.

The materials used included historical databases from a retail company operating in Latin America. These contained weekly demand records ( ) for a period of 24 months, macroeconomic variables such as inflation and exchange rates, and operating conditions such as transport routes, delivery times, and inventory levels. Relevant exogenous data such as weather forecasts, seasonality, and promotional events were also integrated. The tools used for data manipulation and analysis were Python 3.10, with the use of specialized libraries such as Pandas, NumPy, scikit-learn, TensorFlow, and SHAP for model interpretation.

The research design was structured in four phases: data collection and cleaning; development and implementation of prediction models; validation of results using statistical techniques; and simulation of logistical impact. In the first phase, data cleaning techniques were applied to handle missing values, detect outliers, and standardize variables. In the second phase, three AI models (RNN, Transformer, and Gradient Boosting) and two traditional models (ARIMA and exponential smoothing) were trained. The implementation followed a cross-validation strategy with k-fold partitioning (k=5) to ensure the robustness of the results.

In the third phase, the accuracy of the models was evaluated using widely accepted metrics in time series prediction: Mean Squared Error (MSE) and Mean Absolute

Percentage Error (MAPE). The results obtained were subjected to statistical significance analysis, using one-way ANOVA tests to identify differences between models, and Tukey HSD post-hoc tests to determine which models showed statistically significant differences. A confidence level of 95% (α = 0.05) was assumed in all analyses.
In the final phase, logistics scenarios were simulated using the predictions generated by each model to manage inventories and plan distribution routes. Key performance indicators such as storage costs, stockout frequency, overstocking, empty kilometers traveled, and estimated CO_2 emissions were measured. To interpret the contribution of each variable to the model's decisions, Shapley additive explanations (SHAP) were used, which allowed the relative influence of factors such as seasonality, promotions, weather, and historical demand behavior to be evaluated.
This methodological approach made it possible not only to evaluate the predictive capacity of AI models compared to traditional techniques, but also their practical applicability in real logistics contexts. The rigor of the statistical analysis and the replicability of the experiment strengthen the validity of the findings and their potential transferability to other industrial sectors or regions with similar characteristics.

RESULTS

Comparison of accuracy between AI models and traditional methods

Three AI models were trained as RNN, Transformer, and Gradient Boosting and were compared with traditional ARIMA and Exponential Smoothing methods using historical data from a retail chain. The evaluation was performed using cross-validation (k=5) together with the MSE (Mean Squared Error) and MAPE (Mean Absolute Percentage Error) error metrics.

Table 1. *Model accuracy in weekly demand forecasting*

Model	MSE	MAPE (%)
RNN	12.	4.
Transformer	10.2	3.9
Gradient Boosting	11.7	4.5
ARIMA	18.6	7.2
Exponential smoothing	20.1	8.1

All AI models outperformed the results obtained with traditional methods, with Transformers achieving an MSE of 10.2 and MAPE of 3.9%, making it the best by far.

ANOVA corroborated significant differences ($F=24.7$, $p<0.001$), and subsequent Tukey HSD tests confirmed that Transformer was significantly better than ARIMA ($p=0.002$) and Exponential Smoothing ($p=0.001$). These improvements were attributable to Transformer's ability to perform time series forecasting by capturing long-term dependencies as described by Wen et al. (2022).

These findings support the claims of Rocco et al. (2024) that machine learning is superior in situations of greater volatility. Additionally, they are consistent with the Walmart case (Wen et al., 2022) where error was reduced to 3% thanks to Transformers. Conversely, they contrast with Jones (2025) who reported a 20% increase in accuracy with hybrid models and attributed the advantage of AI to the quality of the data used.

Impact of AI on inventory cost reduction

Inventory management was simulated in a distribution center using AI with Transformer architecture and also using historical methods. Storage, stock-out, and overstocking costs were measured over a 6-month period.

Table 2. *Comparative costs (in thousands of USD)*

Method	Storage Cost	Stock Breakage Cost	Total Cost
IA (Transformer)	45.	8.	5
Historical Method	62.7	15.4	78.1

The total cost savings from the AI model was 31.5% (t-test, $p=0.008$), largely due to lower overstocking (SHAP values: promotions = 0.32, seasonality = 0.28). This validates the use of AI for Amazon-style real-time inventory adjustments (Caso, 2024). A regression analysis reported that the model's accuracy explained 68% of the variance in costs ($R^2=0.68$).

The findings support Caso's (2024) claim of 95% accuracy in RNNs, although a Transformer was used here. The difference in percentage could be attributed to heterogeneity between sectors, such as retail and logistics. Furthermore, these findings reinforce the argument of Samun et al. (2025) regarding the synergistic effect of AI optimization, given that the savings exceeded the 25% estimated by Nestlé.

Fleet route optimization using AI

A Deep Learning LSTM model was applied to GPS data from a transport fleet. Demand and weather conditions were integrated. The kilometers traveled empty and CO_2 emissions were measured before and after implementation.

Table 3. *Logistics efficiency pre/post AI*

Metric	Before AI	After AI	Reduction (%)
Empty km	1,200	1,020	15
CO_2 emissions (ton)	4.8	4.1	14.6

The 15% reduction in empty kilometers (Wilcoxon p=0.012) confirms the findings of Li et al. (2024). Through cluster analysis, it was verified that the model performed routes in the urban environment (SHAP: demand = 0.41, climate = 0.19), demonstrating its robustness against external disturbances. Likewise, emissions decreased linearly with km (r=-0.89, p<0.001)), supporting the AI-sustainability link mentioned by the WHO. These data partially support the work of Ricci (2024), who described FedEx's fleet usage, which was +18%, and therefore the impact was much lower. This can be explained by the scale of the operation. The work of Urzúa-Morales (2020) is also incorporated, which demonstrated that the use of advanced models (LSTM vs Gradient Boosting) provided efficiencies in other geographical clusters.
To ensure the robustness of these results, they were replicated in three independent data sets following reproducibility standards (FAIR principles). Rashid (2023) mentions these biases as a shortcoming that must be addressed in the methodological design, which reinforces the argument for the cleaning strategy that should be applied in future research.

CONCLUSIONS

This study showed that Artificial Intelligence (AI) models, especially those based on Transformer and LSTM architectures, significantly outperform traditional methods in terms of demand forecasting for inventory and fleet management. The results showed a reduction in forecast error of up to 50% (45% vs. 38.1% with traditional methods), resulting in a 31.5% reduction in logistics costs and a 15% reduction in empty kilometers traveled. These findings confirm the conclusions of Rocco et al. (2024) and Wen et al. (2022) on the advantages of AI in rapidly changing environments, although the analyses carried out pointed to a critical dependence on data quality, in line with the warnings of Rashid (2023).
The combination of AI with macroeconomic and climate factors not only helped optimize inventory management, as seen with Amazon (Case, 2024), but also helped reduce logistics costs and CO_2 emissions, which is essential for sustainable logistics. However, some discrepancies were observed between sectors regarding the magnitude of the benefits, which requires contextualizing the adaptability of the models, as proposed by Samun et al. (2025) in their hybrid approach.

Journal of Business and entrepreneurial studies
ISSN: 2576-0971
http://journalbusinesses.com/index.php/revista

Barriers include a lack of classified information and organizational culture, which could explain 60% of reported failures in AI implementations (Rashid, 2023). Future research should analyze the possibility of scaling these models in SMEs or PREGUNTOS. It would also be interesting to analyze their integration with blockchain for traceability, as proposed by Charles et al. (2023). To conclude the analysis, it can be added that AI has a positive impact on logistics, but it is necessary to take into account the need for training, quality organizational data, and adaptation to local realities.

REFERENCES

Avella Castro, J. C., & Mendéz Hernández, Á. J. (2021). Inventory classifier using a convolutional neural network synchronized with a database.

Campos Diez Canseco, Hernán. (2017). Digital transformation in the transportation industry: LATAM analysis.

Caso Matos, C. D. (2024). Algorithmic personalized pricing: Case studies.

Charles, V., Emrouznejad, A., & Gherman, T. (2023). A critical analysis of the integration of blockchain and artificial intelligence for supply chain. *Annals of Operations Research*, *327*(1), 7-47.

Jones, J. (2025). AI-driven demand forecasting in supply chains: A qualitative analysis of adoption and impact. https://doi.org/10.20944/preprints202501.1349.v1

Li, A., Zhuang, S., Yang, T., Lu, W., & Xu, J. (2024). Optimization of logistics cargo tracking and transportation efficiency based on data science deep learning models.

Quindimil, M. (2024). Artificial Intelligence and International Trade from a Latin American Perspective. *Global Trade and Customs Journal*, *19*(11/12).

Rashid, H. (2023). Investigating Barriers and Challenges to Artificial Intelligence (AI) Implementation in Logistic Operations: A Systematic Review of Literature. In *Proceedings of the International Conference on Industrial Engineering and Operations Management*.

Ricci, A. (2024). Reinforcement Learning for a Routing Optimization Problem. Solving a VRP with a FedEx data set.

Rocco di Torrepadula, F., Napolitano, E. V., Di Martino, S., & Mazzocca, N. (2024). Machine learning for public transportation demand prediction: A systematic

ISSN: 2576-0971

http://journalbusinesses.com/index.php/revista

literature review. *Engineering Applications of Artificial Intelligence*, *137*, 109166. https://doi.org/10.1016/j.engappai.2024.109166

Samun, T. I., Mahmud, S., Rahman, M., & Ridoy, M. M. H. (2025). A hybrid distribution network model for cost, quality and distribution optimization in perishable goods supply chains. *Malaysian Journal of Science and Advanced Technology*, 1-11.

Urzúa-Morales, J. G., Sepulveda-Rojas, J. P., Alfaro, M., Fuertes, G., Ternero, R., & Vargas, M. (2020). Logistic modeling of the last mile: Case study Santiago, Chile. *Sustainability*, *12*(2), 648.

van Hoek, R. I., & Chong, I. (2001). Epilogue: UPS logistics–practical approaches to the e- supply chain. *International Journal of Physical Distribution & Logistics Management*, *31*(6), 463-468.

Wang, L., & Hsu, H. H. (2025). IoT technology in maritime logistics management: exploration of data analysis methods. *Discover Internet of Things*, *5*(1), 66.

Wen, Q., Zhou, T., Zhang, C., Chen, W., Ma, Z., Yan, J., & Sun, L. (2022). Transformers in time series: A survey. *arXiv preprint arXiv:2202.07125*. https://doi.org/10.48550/arXiv.2202.07125