IBIMA Publishing
Communications of International Proceedings
https://ibimapublishing.com/p-articles/45AI/2025/4548525/
Vol. 2025 (20), Article ID 4548525

Explainable AI in Multi-Warehouse E-commerce Order Fulfilment Optimization[1]*

Piotr POTIOPA

AGH University of Krakow, Faculty of Management
ul. Gramatyka 10, 30-067 Kraków, Poland
ORCID: 0000-0002-8176-7443

Correspondence should be addressed to: Piotr POTIOPA, ppotiopa@agh.edu.pl

* Presented at the 45th IBIMA International Conference, 25-26 June 2025, Cordoba, Spain

Abstract

The rapid expansion of the e-commerce sector has led to complex multi-warehouse networks, where optimizing the order allocation process is a critical operational challenge. While Machine Learning (ML) offers powerful predictive solutions, its widespread adoption in logistics is significantly hindered by the "black box" problem. The lack of model transparency erodes managerial trust, creating a crucial void in the literature for practical, trust-enhancing applications of Explainable Artificial Intelligence (XAI) in supply chain management. To address this gap, this paper proposes and validates a decision support system that integrates a high-performance XGBoost model with the SHAP (SHapley Additive exPlanations) interpretation method. The system was trained and evaluated on a real-world dataset spanning 24 months and over 50,000 orders. Our findings are twofold: first, the model achieved an accuracy of 72.9%, significantly outperforming a baseline heuristic (67.1%). Second, and more importantly, the SHAP analysis demystified the model's decision-making logic, revealing that it prioritizes order complexity and financial value (features *min_font_size, all_products_value*) over mere geographical proximity. This work demonstrates that combining high predictive performance with transparency creates a trustworthy tool capable of overcoming key barriers to AI adoption in logistics.

Keywords: Explainable Artificial Intelligence, XAI, e-commerce, logistics, order fulfilment, XGBoost, SHAP, optimization, supply chain management

Introduction and aim of the work

The e-commerce sector is experiencing unprecedented, dynamic growth; it is estimated that in 2024, global online sales reached a value of approximately 6 trillion USD (Statista, 2024). As the market expands, so do customer expectations, for whom the speed and reliability of delivery have become key factors in choosing a retailer. To meet these demands, many companies are implementing a multi-warehouse strategy, locating distribution centres in various regions to shorten the so-called "last mile" and reduce delivery times (Li and Gao, 2023). Managing a multi-warehouse network, while strategically advantageous, generates a decision-making problem of high complexity. The optimal allocation of each order to a specific warehouse requires the simultaneous consideration of multiple, often competing, factors: customer location, current inventory levels, transportation costs, expected fulfilment time, and the risk of splitting an order into multiple packages, which

[1] The APC was funded under subvention funds for the Faculty of Management and by program "Excellence Initiative - Research University" for the AGH Universi (Adadi and Berrada, 2018) ty of Krakow

Cite this Article as: Piotr POTIOPA, Vol. 2025 (20) "Explainable AI in Multi-Warehouse E-commerce Order Fulfilment Optimization " Communications of International Proceedings, Vol. 2025 (20), Article ID 4548525, https://doi.org/10.5171/2025.4548525

significantly increases costs and complicates the process (de Koster et al., 2007). This issue, known as the order allocation problem, is a complex optimization challenge. In response to this challenge, various methods have been proposed in the literature, ranging from classic mathematical programming models to advanced heuristics and metaheuristics (Nanda and Patnaik, 2023; Hao et al., 2024). Currently, with the growing availability of data, Artificial Intelligence (AI) and Machine Learning (ML) methods are gaining increasing popularity. ML models, trained on historical data, can effectively learn complex, non-linear relationships and automate decision-making processes, which has been successfully implemented by e-commerce giants such as Amazon and Walmart (Cleverence Team, 2024). However, despite promising results, the widespread adoption of advanced AI models in logistics faces a fundamental barrier: a lack of transparency. Many of the most effective algorithms, such as deep neural networks or boosting models, operate as a black box, meaning their decision-making mechanism is incomprehensible to humans (Adadi and Berrada, 2018). As numerous studies indicate, the lack of explainability is one of the main obstacles to implementing AI in supply chain management, as it undermines the trust of managers who bear the ultimate responsibility for operational decisions (Baryannis et al., 2019). This problem concerns not only a lack of trust but also the decision-maker's loss of a sense of control and agency—a manager cannot be fully responsible for a decision whose logic they cannot understand, question, or verify. In response to this problem, the field of Explainable Artificial Intelligence (XAI) has emerged, aimed at creating AI models that can justifiably explain their predictions and recommendations in an understandable way (Arrieta et al., 2020). The integration of XAI with decision-making systems is crucial for building trust and facilitating collaboration between humans and machines, which is a necessary condition for the responsible implementation of AI in critical business areas (Akubilla et al., 2025). Nevertheless, the application of XAI in logistics and supply chain management is still in its early stages of development and constitutes an active area of research (Kosasih et al., 2024; Ordibazar et al., 2025). The aim of this paper is to design, implement, and validate an XAI-based decision support system intended for optimizing the order fulfilment process in a multi-warehouse architecture. The proposed solution integrates a high-performance predictive model (XGBoost) with a modern explanation technique (SHAP) to provide a tool that is not only operationally effective but also fully transparent. The main contribution of this work is to empirically demonstrate that such an approach allows not only for the improvement of key logistics indicators (KPIs) but also for overcoming the barrier of distrust, offering managers a practical path for implementing advanced analytics in daily operations. The remainder of this paper presents a literature review, a detailed description of the research methodology, an analysis of the obtained results, and a discussion of their practical implications.

State of the art

Bloc To precisely position this research within the scientific context, a literature review was conducted covering three key areas: order fulfilment optimization in multi-warehouse networks, the application of Explainable Artificial Intelligence in supply chain management, and the specifics of the XGBoost models and the SHAP method.

A. Order Fulfillment Optimization in Multi-Warehouse Networks

The problem of optimal order allocation has long been a subject of interest for researchers in the fields of logistics and operations research. Early works focused on mathematical programming models, such as integer programming, which aim to find a globally optimal solution but are often impractical for large-scale applications due to their high computational complexity. In response to these limitations, numerous heuristic and metaheuristic methods have been developed. For example, Hao et al. (Hao et al., 2024) applied the NSGA-II evolutionary algorithm to solve a multi-objective model for a multi-warehouse system, while Nanda and Patnaik (Nanda and Patnaik, 2023) used a multi-agent approach to generate order fulfilment scenarios. In recent years, with the Big Data revolution, increasing attention has been given to machine learning-based approaches. These methods utilize historical data for demand forecasting, inventory level optimization, or warehouse operations management (Baryannis et al., 2019). Research shows that ML algorithms can effectively identify complex patterns not captured by traditional models, leading to significant improvements in forecast accuracy and cost reduction. However, as noted in many literature reviews, most of these studies focus on maximizing predictive performance, often neglecting the issue of model interpretability and trust, which remains a key challenge in practical implementations (Kosasih et al., 2024)v.

B. Explainable Artificial Intelligence (XAI) in Supply Chain Management

The "black box" problem of AI has led to the emergence of the XAI field, which aims to "look inside" models and provide understandable explanations of their behaviour (Adadi and Berrada, 2018). In the context of Supply

Chain Management (SCM), XAI is seen as a key enabler for building trust in AI systems and promoting their wider adoption (Baryannis et al., 2019). Systematic literature reviews indicate that this area is still under-researched, and that many existing applications are conceptual or prototypical in nature, with few implementations in real-world industrial settings (Kosasih et al., 2024). Researchers emphasize that the lack of interpretability is a significant research gap that hinders progress in the application of AI in SCM (Baryannis et al., 2019). The need for explanations is particularly acute in decentralized networks, where decisions made at one link in the supply chain affect others, and a lack of understanding of the reasons for these decisions can lead to conflicts and inefficiencies (Ordibazar et al., 2025). XAI methods can be broadly divided into global (explaining the overall behaviour of the model) and local (explaining a single prediction), as well as model-specific (tailored to a specific architecture) and model-agnostic (universal) (Adadi and Berrada, 2018). The choice of the appropriate technique depends on the specifics of the problem, data, and model.

C. Gradient Boosting Models (XGBoost) and the SHAP Method

This work utilizes a combination of the XGBoost model and the SHAP method, a choice motivated by their complementary properties and high effectiveness. XGBoost (eXtreme Gradient Boosting) is an advanced machine learning algorithm based on gradient-boosted decision trees (Chen and Guestrin, 2016). It is widely recognized as one of the most effective models for tabular data, regularly winning competitions in the field of data science. Its strength lies in the combination of speed, scalability, and high predictive accuracy. XGBoost introduces a series of improvements over traditional gradient boosting algorithms, such as regularization to prevent overfitting, native handling of missing data, and parallel processing, making it an ideal choice for complex optimization problems in logistics. SHAP (SHapley Additive exPlanations) is a unified method for explaining the predictions of any machine learning model, introduced by Lundberg and Lee (Lundberg and Lee, 2017). Its theoretical foundation is the concept of Shapley values from cooperative game theory, which fairly attributes the contribution of each "player" (in this case, the model's features) to the final outcome of the "game" (the prediction). SHAP has desirable properties, such as local accuracy and consistency, which means that the explanations are faithful to the model and consistent with human intuition. Importantly, for tree-based models like XGBoost, there exists a highly optimized, model-specific algorithm, TreeExplainer, which calculates SHAP values exactly and efficiently, rather than approximately. This technical synergy makes the combination of XGBoost and SHAP not just a pairing of two powerful tools, but a coherent and methodologically sound framework that maximizes both predictive performance and explanation fidelity.

Research Methodology

The research process was divided into four main stages: (A) data preparation and analysis, (B) predictive model design and training, (C) implementation of the XAI explanation module, and (D) experimental design and definition of evaluation metrics.

D. Dataset and Feature Engineering

Certainly The research was based on a historical dataset from an e-commerce company (from the years 2023-2024) that sells personalized engraved products and fulfils orders from four strategically located warehouses: A (Masovian Voivodeship), D (Lower Silesian Voivodeship), L (Lesser Poland Voivodeship), and P (Lesser Poland Voivodeship). The dataset comprised complete information on over 50,000 orders spanning a 24-month period. Initial data preprocessing consisted of cleaning, deduplication, and mapping customer addresses to geographical coordinates. Subsequently, for each order and each of the four potential warehouses, a set of input features was created. This process included both data directly available in the system and features generated through feature engineering. The key input features, detailed in Table I, included order characteristics, customer data, product information, and relational features such as distance and inventory availability.

E. Predictive Model Architecture

The order allocation problem was formulated as a multiclass classification task. The model's objective is, for a given order, to predict which of the four warehouses (A, D, L or P) is the optimal location for its fulfilment. The target variable (label) in the training process was the warehouse from which a given order was actually fulfilled in the historical data. The assumption was made that historical decisions, while not always perfect, provide a sufficiently good approximation of optimality to enable effective model training in a supervised learning mode. The entire dataset, comprising over 50,000 orders, was chronologically split into a training set (80% of the data, over 40,000 orders) and a test set (20%, approx. 10,000 orders). This split simulates a real-world scenario where the model is tested on "future," previously unseen order data. The XGBoost algorithm was used to build the predictive model. Its hyperparameters were optimized on the training set using a grid search (GridSearch) method with 5-fold stratified cross-validation. To further counteract the class imbalance in the data, a sample weighting mechanism was used during the training process, which ensured the model's robustness and minimized the risk of overfitting. The optimization goal was also the maximization of the weighted F1-score metric, which is a balanced measure of precision and recall, particularly useful in the case of imbalanced classes. The most important optimized hyperparameters are presented in Table II.

F. Implementation of the XAI Explainer Module

After training the final XGBoost model, an explainer module was implemented using the shap library in Python. The *shap.TreeExplainer* mechanism was used, which is optimized for tree-based models and allows for the exact calculation of SHAP values for each feature and each prediction. For each order in the test set, in addition to the warehouse prediction, a full set of SHAP explanations is generated. These explanations are then used to create visualizations that form the basis of the analysis in the latter part of this paper:

- Global Feature Importance Plots (Summary Plots): These show which features have the greatest average impact on the model's decisions across the entire dataset (Fig. 1).
- Local Decision Plots: These illustrate how individual features contribute to the final prediction for a single, specific order, showing the model's "decision path" (Fig. 2).

This approach enables the model to be analysed on two levels: strategic (understanding the overall operational logic) and operational (understanding the reasons for a specific recommendation).

G. *Experimental Design and Evaluation Metrics*

To evaluate the effectiveness of the proposed solution, a simulation was conducted on the test set. The results of the XGBoost model with the XAI module were compared with two baseline strategies:

1. **Historical Strategy:** The actual order fulfilment as recorded in the data. This serves as a reference point for assessing data quality and the potential for optimization.
2. **Baseline Heuristic:** A simple, intuitive decision rule: "assign the order to the geographically closest warehouse that has 100% of the required products in stock." If no warehouse meets this condition, the order awaits stock replenishment.

The evaluation was conducted using two groups of metrics:

- **Predictive Performance Metrics:** Accuracy, Precision, Recall, and F1-score, to assess how well the model replicates (and potentially improves upon) historical decisions.
- **Business Metrics (KPIs):** Average delivery time (in hours), average shipping cost (in PLN), and the percentage of back-ordered orders due to stock shortages at a given warehouse.

Results and Analysis

This section presents the results of the experiments, covering both a quantitative evaluation of the model's effectiveness and a qualitative analysis of its interpretability using XAI techniques

A. Evaluation of the Model's Predictive Performance

In the first step, the ability of the XGBoost model to predict the optimal warehouse was evaluated. Table III presents the classification results on the test set in comparison to the simple baseline heuristic. To evaluate the predictive performance of the developed XGBoost model, its results on the test set were compared with those of a simple baseline heuristic. A comparison of key performance metrics is presented in Table III. The data analysis clearly indicates a significant advantage of the machine learning model over the heuristic approach. The XGBoost model achieved an accuracy of 72.9%, which represents a distinct improvement of 5.8 percentage points over the heuristic (67.1%). This means that the model makes significantly more correct decisions regarding the choice of the optimal warehouse, which directly translates into a better alignment with actual logistical needs. The model's superiority is evident across all analysed metrics. The weighted precision increased from 73.4% to 78.0%, indicating that the model's predictions are more reliable, resulting in fewer incorrect assignments to the wrong warehouse. Simultaneously, the increase in weighted recall from 68.1% to 73.1% shows that the model is able to more effectively identify all instances belonging to a given class. Finally, the F1-Score, being a harmonic mean of precision and recall, also confirms the overall superiority of the model, reaching a value of 74.3% compared to 69.9% for the heuristic. These results demonstrate that the XGBoost model has successfully learned complex, non-linear relationships in the data that go beyond the simple business logic implemented in the heuristic. This provides a solid basis for concluding that the model's implementation can bring tangible benefits to the logistics optimization process.

Table I: Description of Input Features Used in the Predictive Model

Category	Feature Name	Description and Role
Order Characteristics	*payment_method*	Categorical payment method (e.g., bank transfer, PayU). After one-hot encoding, it becomes a set of binary features.
	shipping_method	Categorical shipping method (e.g., InPost courier, personal pickup). Also subject to one-hot encoding.
	all_products_value	Total order value including shipping costs.
	all_positions	Total number of unique line items (products) in the order.
	all_products_cnt	Total count of all product units in the order (sum of quantities).
	all_products_sum	Total value of the products alone (excluding shipping costs).
	latitude	Geographical latitude of the delivery address. Used for distance calculation.
	longitude	Geographical longitude of the delivery address. Used for distance calculation.
	order_date	Date and time the order was placed. Can be used for seasonality analysis.
	release_date	Planned or actual date the order was shipped from the warehouse.
	user_ip	Customer's IP address, potentially used for geolocation.
	zipcode	Postal code of the delivery address.
	city	City of the delivery address.
	all_fonts	Set of unique fonts used in the projects within the order.
Product Characteristics	*min_font_size*	Minimum font size used in the projects within the order. May indicate complexity.
	max_font_size	Maximum font size used in the projects.
	flash_technology	Binary variable (1/0) indicating if the order contains a product made with "flash" technology.
	num_of_categories	Number of unique product categories in the order.
	case_colors	Number of unique case colours in the order.
Relational Features (Warehouse-Specific)	*X_all_in_stock*	Binary variable (1/0) indicating whether warehouse X has all products from the order in stock. A key decision factor.
	dist_X	Haversine distance (in km) from the customer's delivery address to warehouse X.
	dist_if_avail_X	Conditional feature: the distance to warehouse X if it has all products in stock. Otherwise, it assumes a very large value (99999), strongly "penalizing" the model for choosing an unavailable option.

Table II: Optimal Hyperparameters for the XGBoost Model Found by GridSearch

Hyperparameter	Optimal Value
n_estimators (number of trees)	**200**
max_depth (maximum tree depth)	**7**
learning_rate (learning rate)	**0.1**
subsample (fraction of data per tree)	**0.8**

Table III: Predictive Performance Comparison of the Models

Metric	Baseline Heuristic	XGBoost Model
Accuracy	67.1 %	**72.9 %**
Precision (weighted)	73.4 %	**78.0 %**
Recall (weighted)	68.1 %	**73.1 %**
F1-Score (weighted)	69.9 %	**74.3 %**

B. Methodology for Assessing Impact on Key Performance Indicators (KPIs) and Future Research Directions

After confirming the predictive effectiveness of the XGBoost model, the next key step is to assess its real-world impact on business and operational indicators. As part of future work, a detailed simulation analysis is planned to quantify the benefits of implementing the developed warehouse prediction mechanism. The aim of this analysis will be to compare three operational scenarios for the historical set of orders:

1. **The Historical Scenario** – based on the actual, archived decisions.
2. **The Heuristic Scenario** – in which the warehouse is assigned based on a simple business rule (e.g., the geographically closest).
3. **The Predictive Scenario** – in which the warehouse is selected based on the recommendation of the XGBoost model.

The comparative analysis of these scenarios will focus on Key Performance Indicators (KPIs) that reflect both the cost and quality aspects of the logistics process. The key measured indicators will include:

- **Average end-to-end order fulfilment time** – measured from the moment a customer places an order until its delivery. This is a key indicator influencing customer satisfaction and loyalty.
- **Average shipping cost** – directly reflecting the financial savings generated by the model.
- **Percentage of split orders** – an indicator whose minimization translates into lower operational costs and a better customer experience.

Conducting the aforementioned analysis will be of fundamental importance. Streamlining the order fulfilment process through more accurate predictive decisions will create the opportunity to thoroughly examine the impact of these decisions on both qualitative benefits (e.g., increased customer satisfaction, shorter delivery times) and cost savings (e.g., reduced transportation and handling costs). The results of this research will form the basis for business recommendations regarding the full-scale implementation of the system in a production environment and will serve as a starting point for further supply chain optimization.

C. Global Model Interpretation using SHAP

To understand the principles on which the model makes such effective decisions, a global interpretability analysis was conducted using SHAP. Figure 1 presents a feature importance plot, ranking the features by their average impact on the model's predictions. Analysis of the SHAP feature importance plot allows for a deep understanding of the model's decision-making strategy. Contrary to the initial assumptions of the baseline model, the most important factors are not directly distance or general availability, but rather features that precisely describe the specifics of the order itself. At the top of the ranking, with the greatest average impact on predictions, are features such as *num_min_font_size*, *num_all_products_value*, and *num_all_products_sum*. This indicates that the model has learned that the financial value of the order and its complexity are of key importance for warehouse selection, potentially represented by the unusual feature *min_font_size*, which may be a significant indicator of personalized products. Another group of significant factors are the variables concerning inventory availability in individual warehouses, such as *num_A_all_in_stock* or *num_L_all_in_stock*. Interestingly, the plot shows that availability in one warehouse (e.g., A) has a strong influence on the probability of selecting another warehouse (e.g., L), which proves that the model has learned complex relationships and trade-offs between the locations. Only further down the ranking do features related to payment methods (*cat_payment_method_bank* transfer) and the distance to individual warehouses (e.g., *num_dist_if_avail_L*) appear. This global perspective is extremely valuable. It reveals that the AI's optimization strategy is far more advanced than a simple heuristic. The model intelligently prioritizes orders based on their value and characteristics, only then considering geographical and operational factors. This allows managers to understand that the key to optimization in this case is not simply reducing distance, but intelligently managing the flow of orders with diverse characteristics.

D. Local Decision Interpretation – A Case Study

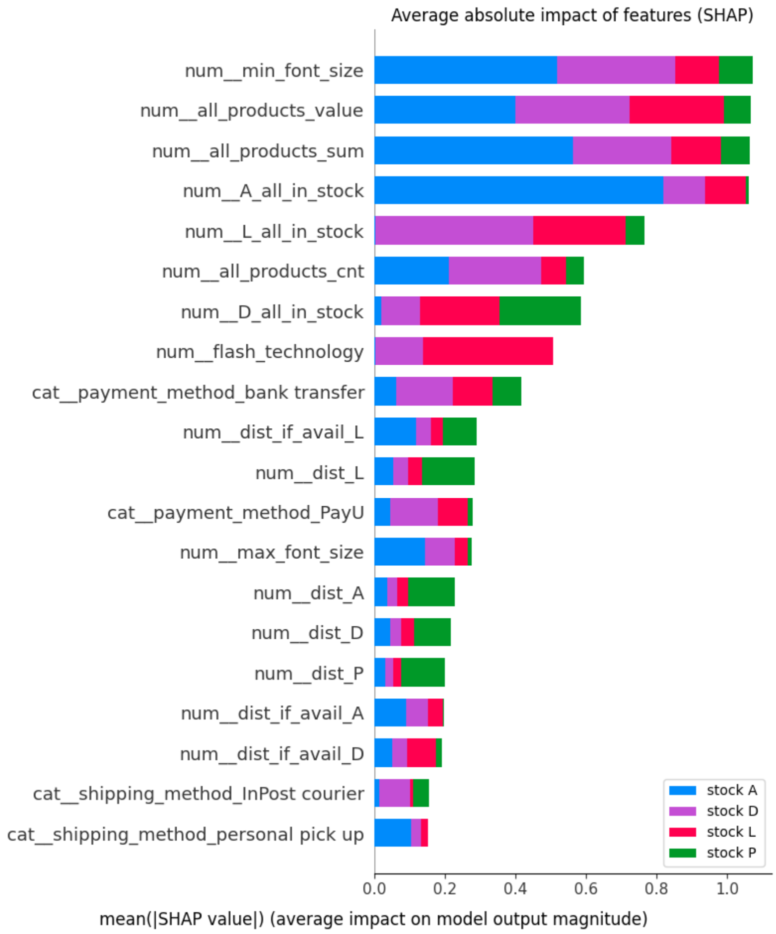

Figure 1. Global feature importance according to SHAP (example plot).

The greatest strength of XAI, however, is the ability to analyse single, specific decisions. Let us consider the case of order #1C917, for which the model made a decision seemingly contrary to simple logic. A customer from Marki (near Warsaw) ordered 1 product. The nearest warehouse, L (Masovian Voivodeship), was located approx. 130 km away, while warehouse A (Lesser Poland Voivodeship) was approx. 330 km away. The baseline heuristic would have chosen warehouse L. However, the XGBoost model recommended shipping from the more distant warehouse A, primarily due to product availability, but also considering the simpler project design (which can be handled at that facility) and the small order size. Figure 2 presents the SHAP explanation for this decision. The SHAP decision plot clearly shows the model's logic. The base value (the average prediction) is pushed towards the final decision by the contributions of individual features. The strongest factor pushing the prediction towards selecting warehouse A is *min_font_size* (SHAP value +0.73). This indicates that warehouse A handles specific types of projects (engraved products). Conversely, the proximity to warehouse L (represented by the *distance_L* feature) had a negative impact on the choice of warehouse A (SHAP value -0.1597).

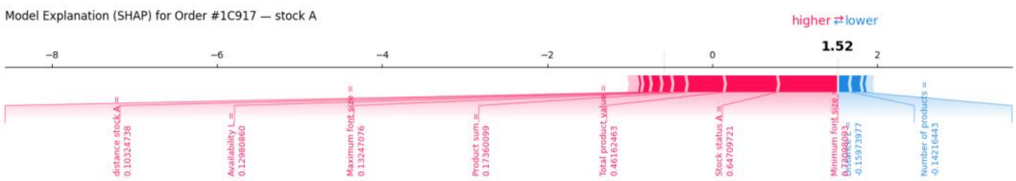

Figure 2. SHAP explanation for the warehouse A recommendation for order #1C917.

However, the positive impact of availability at warehouse A (*Stock_status_A*) was much stronger (+0.64). The model determined that it was better to ship the order from a single, more distant location than to handle it at a busier warehouse that may be responsible for fulfilling more complex orders. Such transparent reasoning allows a manager not only to accept the AI's recommendation but also to understand its logic and build trust in the system.

Discussion and Practical Implications

The obtained empirical results lead to several important conclusions of both a technical and a business nature. The combination of a high-performance predictive model with an advanced explanation technique allows for the creation of a system that is not only effective but also trustworthy. The main achievement of the presented approach is the demonstration that it is possible to simultaneously achieve two, seemingly conflicting, goals: maximizing operational efficiency and ensuring full decision-making transparency. The XGBoost model can significantly improve key logistics indicators, which directly translates into cost reduction and increased customer satisfaction—this will be the subject of further analysis and research in this e-commerce area. Simultaneously, the SHAP module demystifies the decision-making process by answering the fundamental question, "why?". This duality is the key to overcoming AI adoption barriers in e-commerce logistics.

The practical implications of this solution for logistics managers are far-reaching. The proposed system is not a "black box" that imposes decisions, but rather an intelligent assistant that supports the decision-making process on several levels:

- **Decision Support:** The system provides a real-time recommendation for the optimal warehouse for each new order.
- **Auditing and Trust Building:** A manager can, at any time, analyse the SHAP explanation for any recommendation to understand its basis. This allows for the verification of the AI's logic and helps build trust in the system.
- **Exception Handling:** In the case of recommendations that are contrary to intuition (as in the analysed case of order #1C917), the SHAP explanation provides a rational justification. The manager can either accept the new, nonobvious strategy discovered by the AI or, armed with the knowledge of what the model considered, consciously make a different decision by leveraging their domain expertise.
- **Knowledge Discovery:** Analysis of the global SHAP explanations (Fig. 1) can reveal systemic, previously unknown relationships within the logistics network. This can lead to strategic conclusions, e.g., regarding the stocking policies of individual warehouses or contract negotiations with carriers.

The implementation of such a system enables the transformation of logistics management from a reactive model (e.g., firefighting, resolving delays and customer complaints) to a proactive model. In line with the reasoning presented in the literature (Arrieta et al., 2020), managers can use the system's predictions and recommendations to take pre-emptive actions that prevent problems from arising, rather than merely reacting to their consequences. This research directly addresses the research gaps identified in recent literature reviews, which call for more application-oriented research on XAI in SCM and for the design of systems that take a holistic approach to decision-making problems and consider the need for interpretability at every stage (Aljohani, 2023).

Conclusions and Future Research Directions

This paper has demonstrated that the integration of an XGBoost machine learning model with the SHAP explainer mechanism constitutes an effective and transparent solution to the order allocation optimization problem in a multi- warehouse e-commerce network. The proposed system can not only significantly improve key logistics performance indicators, such as delivery time and cost, and the percentage of delayed orders (which will be the subject of further in- depth research), but more importantly, it makes the AI's decision-making process fully understandable to the end-user. This transparency is crucial for building trust and facilitating the practical implementation of advanced AI technologies in business operations.

One of the key advantages of the developed model is its ability to adapt to dynamic operational conditions. Unlike standard approaches, the model was trained using data on the current stock levels for each product included in an order, which allows it to consider inventory availability in near real-time. Nevertheless, to fully realize the potential of a flexible logistics system, the next research step should be to enrich the model with data concerning the current production and operational load of individual warehouses. The integration of production planning with distribution is cited as a key factor in increasing the efficiency and resilience of modern supply chains (Chen and Vairaktarakis, 2005). Introducing such a feature would allow the model not only to select a warehouse with available goods but also to prefer those locations that currently have spare operational capacity, which would directly translate into a reduction of the total order fulfilment time. It should also be noted that the current evaluation was conducted based on simulations and historical data. The ultimate confirmation of the model's value would be to conduct A/B tests in a live production environment, which would allow for a direct measurement of its impact on key performance indicators. These two areas—integration with operational load data and deployment testing—define the main directions for future work.

1. **Integration with Real-Time Systems:** Expanding the model to allow for integration with Warehouse Management Systems (WMS) to consider current, rather than historical, inventory levels and operational load.

2. **Interactive Decision Dashboards:** Creating a user-friendly graphical interface to present recommendations to managers along with interactive SHAP visualizations. This would enable intuitive data exploration and drill-down into the model's logic, which is crucial for its practical usability (Sudjianto and Zhang, 2021).

3. **Application of Reinforcement Learning (RL):** Investigating the use of RL algorithms where an agent would learn an optimal allocation policy not just from historical data, but by interacting with the environment and receiving rewards for good decisions (e.g., low cost, short delivery time).

4. **Holistic Modeling:** Expanding the model's scope to include upstream factors (e.g., supplier reliability, ware- house inbound lead times) and downstream factors (e.g., availability and capacity of courier companies) to create a more comprehensive, holistic decision support system for the supply chain.

In conclusion, the work presented here represents a step towards building more intelligent, efficient, and, most importantly, trustworthy logistics systems. It shows that artificial intelligence does not have to be a "black box," but can instead become a transparent partner in the decision-making process.

References

- Adadi, A. and Berrada, M. (2018) 'Peeking inside the black-box: a survey on explainable artificial intelligence (XAI),' IEEE Access, 6, 52138–52160. DOI: 10.1109/ACCESS.2018.2870052.
- Akubilla, J., Somoye, O. I., Abiodun, F. and Serifat, O. A. (2025) 'The role of explainable AI in enhancing trust and transparency in supply chain risk mitigation'. URL: https://www.allmultidisciplinaryjournal.com/uploads/archives/20250509150013_MGE-2025-3-065.1.pdf
- Aljohani, A. (2023) 'Predictive analytics and machine learning for real-time supply chain risk mitigation and agility,' *Sustainability*, 15 (20), 15088. DOI: 10.3390/su152015088
- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R. and Herrera, F. (2020) 'Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI' *Information Fusion*, 58, 82-115. DOI: 10.1016/j.inffus.2019.12.012.
- Baryannis, G., Validi, S., Dani, S. and Antoniou, G. (2019) 'Supply chain risk management and artificial intelligence: state of the art and future research directions,' *International Journal of Production Research*, 57 (7), 2179-2202. DOI: 10.1080/00207543.2018.1530476.
- Chen, T. and Guestrin, C. (2016) 'XGBoost: a scalable tree boosting system,' *in Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. DOI: 10.1145/2939672.2939785.
- Chen, Z-L. and Vairaktarakis, G. L. (2005) 'Integrated scheduling of production and distribution operations,' *Management Science,* 51 (4), 614-628. DOI: 10.1287/mnsc.1040.0325.

- Cleverence Team. (2024) ‘How Amazon and Shopify use AI to optimize order fulfillment and delivery,’ [Online]. Cleverence. [Accessed 20 June 2025]. Available: https://www.cleverence.com/articles/business-blogs/how-amazon-and-shopify-use-ai-to-optimize-order-fulfillment-and-delivery/.
- de Koster, R., Le-Duc, T. and Roodbergen, K. J. (2007) 'Design and control of warehouse order picking: a literature review,' *European Journal of Operational Research,* 182 (2), 481-501. DOI: 10.1016/j.ejor.2006.07.009.
- Hao, L., Xie, H. and Liu, Y. (2024) ‘A multi-objective optimization model for single-product and multi-warehouse based on NSGA-II,’ *in Proc. of 21st Int. Computer Conf. on Wavelet Active Media Technology and Information Processing (ICCWAMTIP),* 1118-1121. DOI: 10.1109/ICCWAMTIP64812.2024.10873689.
- Kosasih, E. E., Papadakis, E., Baryannis, G. and Brintrup, A. (2024) ‘A review of explainable artificial intelligence in supply chain management using neurosymbolic approaches,’ *International Journal of Production Research*, 62 (4), 1510–1540. DOI: 10.1080/00207543.2023.2281663.
- Li, W. and Gao, G. (2023) ‘Research on multi-product order splitting and distribution route optimization of ‘multi-warehouse in one place’,’ *Frontiers in Business, Economics and Management,* 8 (3), 138–142. DOI: 10.54097/fbem.v8i3.7449.
- Lundberg, S. M. and Lee, S.-I. (2017) ‘A unified approach to interpreting model predictions,’ in *Advances in Neural Information Processing Systems 30 (NIPS 2017),* 4765–4774. DOI: 10.48550/arXiv.1705.07874.
- Nanda, P. and Patnaik, S. (2023) ‘A multi-agent coalition-based approach for order fulfilment in e-commerce,’ *Decision Analytics Journal*, 7, 100227. DOI: 10.1016/j.dajour.2023.100227.
- Ordibazar, A. H., Hussain, O. K., Chakrabortty, R. K., Irannezhad, E. and Saberi, M. (2025) 'Artificial intelligence applications for supply chain risk management considering interconnectivity, external events exposures and transparency: a systematic literature review,' *Modern Supply Chain Research and Applications*. DOI: 10.1108/MSCRA-10-2024-0041.
- Statista. (2024) ‘Global retail e-commerce sales 2022–2028,’ [Online]. Statista.com. [Accessed 9 June 2025]. Available: https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/.
- Sudjianto, A. and Zhang, A. (2021) 'Designing inherently interpretable machine learning models,' in *Proc. of ACM ICAIF 2021 Workshop on Explainable AI in Finance*. DOI: 10.48550/arXiv.2111.01743.