SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 4, (2026), pp. 9491-9502
Open Access. Online & Print

DOI: 10.5281/zenodo.12426930

DIGITALIZATION OF SUPPLY CHAINS: CULTURAL AND ORGANIZATIONAL IMPACTS OF AI-DRIVEN OPERATIONS

Dr. Nazia Qureshi[1*], Mona Piyush Rewatkar[2], Dr. Sanjeev Kumar Gautam[3], Dr. Virendra Kumar Verma[4], Ranjan Banerjee[5], Fernando Huamán Ccanto[6]

[1]Assistant Professor, Faculty of Commerce and Management, Mittal Institute of Education, Bhopal, Email ID: nazia.leo@gmail.com*
Orcid ID- 0009-0000-7556-9208
[2]Phd Scholar, Faculty of Commerce and Management, School of Allied Sciences, Datta Meghe Institute of Higher Education and Research Deemed to be University Sawangi (Meghe) Wardha, Email ID: monahiware@gmail.com
Orcid ID- 0009-0006-4164-4269
[3]Associate Professor, Amity University, Noida, Email ID: sanjeevgautam021078@gmail.com
[4]Assistant Professor, Symbiosis Institute of Business Management, Symbiosis International (Deemed University), Bengaluru, India Email ID: virendra.verma@sibm.edu.in Orcid ID: 0000-0002-1115-1221
[5]Assistant Professor, Computer Science & Engineering, Brainware University, Barasat, Kolkata – 125, West Bengal, India 700125, Email ID: rbkpccst@gmail.com Orcid ID: 0009-0003-1950-7530
[6]Universidad Nacional Intercultural de la Selva Central Juan Santos Atahualpa, Perú
Email ID: fhuaman@uniscjsa.edu.pe Orcid ID : 0000-0001-6951-4796

Received: 12/12/2025
Accepted: 30/01/2026

Corresponding Author: Nazia Qureshi
(nazia.leo@gmail.com)

Abstract

The high pace of digitalization of supply chains has placed artificial intelligence (AI) at the core of the processes of operational governance and institutional change. This paper explores the way AI-based predictive systems transform supply chain governance, workforce distribution, and organizational decision standards. Based on archival transactional shipment data, machine learning models under supervision were trained to forecast the risk of late-delivery in a global supply chain environment. Ensemble models were able to discriminate well and were also robust in time. Along with the predictive accuracy, the probability threshold calibration was also operationalized as a governance lever to strike the balance between cost efficiency and service reliability. The cost-sensitive model determined an optimal decision point that minimized the total cost of operation and minimized the number of missed high-risk shipments. Workload simulations converted model results into the daily review volumes and shift-level oversight requirements, which demonstrates the redistribution of organizational attention by AI. A simulation of an intervention also suggested that major shipment categories would have their delays reduced by 31%. The explainability analysis using SHAP was able to identify structural operational drivers of risk and to translate the predictive outputs into policy actions that are governance-oriented. The results indicate that AI-driven digitalization is not limited to performance optimization but it essentially restructures institutional accountability, risk tolerance, and the authority to make decisions in the supply chain activities.

Keywords - Digitalization; Supply Chain Governance; Artificial Intelligence; Predictive Analytics; Algorithmic Decision-Making; Explainable AI; Cost-Sensitive Learning; Organizational Transformation.

1. Introduction

The high rate of digitalization of global supply chains has radically changed the way organizations operate, react to disruptions, and organize complex interdependencies. The recent crises, such as the COVID-19 pandemic, have shown the vulnerability of the interconnected supply chains and the necessity to have resilient and functional systems that can respond to the crisis in an adaptive manner (Ivanov and Dolgui, 2020). In this regard, Industry 4.0 technologies, advanced analytics, and artificial intelligence have become important drivers of digital supply chain initiatives that have become essential enablers of visibility, agility, and strategic coordination (Queiroz et al., 2021). The digital supply chain paradigm shifts the traditional automation to integrated and data-driven ecosystems that can support real-time decision-making and predictive intervention (Buyukozkan and Gocer, 2018). With the growing integration of AI and big data analytics in the workflows of organizations, digitalization is no longer an issue of technological improvement, but it is a structural change in the governance of supply chains. That is well understood that big data analytics and machine learning are major contributors to the development of operational performance and dynamic capabilities (Wamba et al., 2017). In the context of supply chains, predictive analytics are used to assist in forecasting, demand planning, and risk assessment and allow companies to transition to anticipatory control instead of reactive management (Sanders, 2016). Operations management advanced analytics have proven to result in quantifiable efficiency, service reliability, and decision quality improvement (Choi et al., 2018). Moreover, end-to-end digital visibility technologies enable companies to increase resilience and survivability through the use of real-time information streams (Ivanov, 2021). Though previous studies focused on demand forecasting and optimization (Carbonneau et al., 2008), the modern AI system is also more engaged in classification and risk prediction work that directly affects the operational prioritization and intervention plans.

Although the technological advancement of predictive analytics is impressive, a significant gap in comprehending the way AI-driven operations transform organizational governance and cultural standards in the supply chains persists. The development of algorithmic systems to produce risk scores and recommended actions starts to affect the managerial discretion, task allocation, and accountability systems. Algorithms are no longer unbiased instruments; they are the control and coordination mechanisms that transform the dynamics of the workplace (Kellogg et al., 2020). By implementing learning algorithms into the workflow, the decision-making process, the distribution of authority, and the operational priorities are altered (Faraj et al., 2018). In this aspect, the adoption of AI is not only a technical change but a socio-organizational change that influences the decision-making organization (Shrestha et al., 2019). This transformation is complicated by the automation-augmentation paradox, which supports and limits managerial agency at the same time (Raisch and Krakowski, 2021). On the same note, human-AI symbiosis in decision-making demands novel coordination, trust, and institutionalized control (Jarrahi, 2018). It is on this basis that the central issue that this research paper seeks to address is the fact that there is little empirical research that investigates how AI-based predictive systems can be converted into organizational governance mechanisms in the context of supply chains. Although previous research focuses on digital capabilities and performance results, there are less studies in which the predictive thresholds, risk prioritization, and algorithmic explainability reform the workload distribution, intervention intensity, and institutional accountability. Without such integrative analysis, we can only understand digitalization as a structural phenomenon of the organization and not a technological improvement only.

This paper fills that gap by evaluating how a machine learning-based late-delivery risk prediction system can be implemented into a digital supply chain setting. The areas covered by the research include predictive model validation, threshold calibration, cost-sensitive governance analysis, workload redistribution modeling, and explainability-based policy translation. The study conceptualizes AI configuration as a governance tool by balancing technical performance assessment and organizational impact assessment as a method of re-balancing risk tolerance and operational attention. The study is however restricted to transactional supply chain information and does not directly assess employee attitudes or behavioral adaptation. Structural inferences of cultural influences are hence determined structurally by variations in decision logic, workload patterns and authority distribution and not by measurement using surveys.

This study is important because it combines predictive analytics with the organizational and cultural transformation theory. With the growth of digital supply chains being more data-driven, AI systems influence the norms of decisions and accountability mechanisms, in addition to operational efficiency. The study proves the reorganization of supply chain operations through the quantification of workload implications, cost trade-offs, and cross-market stability. This adds to the wider discussion of the digital transformation through the connection between the implementation of AI and institutional adjustment and governance rebalancing.

The aim of the study is thus to empirically research on the impact of AI-based predictive systems on supply chain governance and organizational structures in digital transformation. In particular, the research will address the following questions: (1) how well machine learning models predict operational risk; (2) how well probability thresholds can be used as a governance mechanism in operational risk; (3) how well operational risk can be benefited by AI-guided intervention; and (4) how predictive model explainability can be translated into structured organizational policy. In this integrative manner, the research will put AI in a new role as not just an optimization tool but a driver of organizational and cultural change in digitally enabled supply chains.

2. Literature Review

Digitization of organizational processes has also led to a complete reevaluation of the connection between technology, work, and governance structures. The modern literature is increasingly unconvinced by the traditional distinction between the technological systems and organizational practices and stresses on their intertwined and co-constitutive character. The sociomateriality approach holds that technology and organizational action cannot be practically separated because material objects and human agency constantly influence each other in the context of operations (Orlikowski and Scott, 2008). In digitally enabled supply chains, predictive analytics systems are not only used to support the decisions; they are also engaged in the organization of workflows, setting priorities, and designing accountability systems. Such perspective redefines AI-powered operations as the part of embedded sociomaterial formation as opposed to neutral computational instruments. It is based on this view that studies of affordances and imbrication can further elaborate on the way the digital technologies and organizational routines can be adjusted to each other over time. Leonardi (2011) shows that technologies present constraints and affordances that shape the reorganization of routines by actors. Probability thresholds, dashboards, risk scores in the context of predictive supply chain systems serve as material characteristics of the system that allow some actions (proactive intervention) and restrict others (discretionary decision-making based on intuition only). With the encounter of flexible routines and flexible technologies, new coordination and supervision patterns are introduced, which change the norms in operations and redistribute the decision rights. This engagement highlights the fact that AI implementation is a structural change and not performance maximization.

The reconfiguring of organizational logic by digital technologies is also emphasized by the literature on digital innovation. Yoo et al. (2010) explain that digital architectures can facilitate modularity, generativity, and distributed innovation and transform the way organizations organize complex activities. The predictive analytics systems are embedded in a wider digital infrastructure in a supply chain environment, which incorporates data streams across functioning units. Digital platforms create new organizing logic that makes the decision-making processes more responsive in real-time and mediated by algorithms. With the development of the digital era of service ecosystems, the concept of value creation is increasingly reliant on information flows and analytical strengths instead of physical coordination (Barrett et al., 2015). These theoretical advances imply that AI-based supply chains are not only technological advances but also service and governance formations.

Although the concept of digital transformation has been studied on structural change, the growing use of machine learning presents further concerns in terms of transparency and interpretability. The more complex the predictive model is, the more difficult it is to interpret and justify algorithmic decisions, which is a problem facing organizations. The SHAP values by Lundberg and Lee (2017) are a single model prediction interpretation framework that offers consistent and theoretically motivated explanations of feature contributions. The interpretability methods can help organizations convert statistical results into actionable information, which facilitate institutional trust and integration of governance. To add to this, Molnar (2020) stresses that interpretable machine learning makes it easier to be accountable and to deploy in high-stakes areas. Doshi-Velez and Kim (2017) suggest that to have a rigorous science of explainable AI, interpretability is a necessity, especially when decisions have an impact on organizational performance and interests of stakeholders. All these studies note that, in combination, make explainability not a nice-to-have, but a necessity when integrating AI in governance.

In addition to interpretability, there is a need to have standards of caution in operational deployment of predictive systems to balance decision thresholds and cost trade-offs. Cost-sensitive learning offers a formal model of the use of asymmetric misclassification costs in model design. Elkan (2001) shows that the performance of classification should be measured on the basis of the decision costs and not just accuracy in situations where false negative and false positive have different ramifications on the organization. Equally, Drummond and Holte (2006) suggest that cost curves are a better way of illustrating the performance of a given classifier under different cost assumptions, which supports the need to evaluate the performance of a decision-centric way. Practically speaking, according to Provost and Fawcett (2013), model evaluation should be consistent with business goals, and threshold optimization and cost-benefit arguments are the key elements of data-driven decision-making. Such lessons are especially applicable in supply chain settings whereby missed disruptions can be more expensive than over-reviewing so that positioning threshold selection is a governance tool.

With the rise of algorithmic systems in the organizational processes, the issues of fairness and bias have become prominent. According to Barocas et al. (2019), machine learning systems have the ability to recreate structural inequities, which makes it necessary to conduct systematic assessment of fairness in deployment scenarios. Mehrabi et al. (2021) conduct an extensive review of bias in machine learning, listing the sources of representational and allocative inequalities that could occur throughout data preprocessing, model training, and selection of decision threshold. In global supply chains, where operations cut across geographically different markets, the consideration of fairness is critical to the institutional legitimacy and cross-unit acceptability. The combination of fairness diagnostics and performance assessment is one of the signs of a more general transition to responsible AI governance.

Altogether, this literature provides three pillars on which one can base the understanding of AI-driven supply chain transformation. First, sociomaterial and digital innovation lenses show that technologies transform the organizational structures and decision logics. Second, the interpretability research highlights the importance of transparency in integrating predictive systems into the governance systems. Third, fairness-oriented and cost-sensitive scholarship focuses on the fact that algorithmic configuration is associated with normative and economic trade-offs. Combining these theoretical streams, the current paper places the AI-based supply chain digitalization in a larger context of organizational governance and cultural change and how predictive thresholds, explainability, and performance calibration institutionalize new norms of operation.

3. Methodology

The present study has a design-science and quantitative approach to research because it aims to analyze how AI-based predictive systems redefine organizational governance and cultural frameworks in the context of supply chain management. The methodology is a combination of machine learning under supervision, simulation modeling, and the analysis of the impact of organizations to assess the technical performance and institutional implications of predictive digitalization. Instead of concentrating on the accuracy of the models, the research design incorporates governance-based performance indicators (redistribution of workloads, cost-efficient, fairness of segmentation, and timeliness). By doing so, it allows the empirical evaluation of the impact of the algorithmic systems on the operational structures, the decision rights, and the intervention policies.

3.1 Research Design

The research is an explanatory predictive study. At stage one, predictive models are created to estimate the risk of late delivery in a digitalized supply chain setting. In the second stage, the outputs of the model are converted to governance simulations to assess the implications of workload, cost trade-offs, and policy implications. During the third step, explainable artificial intelligence methods are used to derive meaning of risk drivers and relate them to organizational transformation processes. This design with multiple layers enables the study to connect technical analytics with the institutional and cultural implication, and thus connect predictive modeling with the general theory of digital transformation.

The predictive modeling phase uses comparisons between various supervised learning models such as Logistic Regression, HistGradientBoosting and XGBoost classifiers. The model comparison provides the strength and protects against the bias of algorithms. Generalizability is evaluated using cross-validation procedures and time-based validation simulates real-world deployment conditions and determines the time stability. This mix of the static and time validation enhances the internal validity of the results.

3.2 Data Collection Methods

The empirical analysis is founded on the archival secondary data, which is the publicly available DataCo Smart Supply Chain for Big Data Analysis dataset located on Kaggle (Shashwatwork, 2019).

The data set includes around 180,000 order-level transactional data on world supply chain operations in various geographic markets. Every record is linked to a single shipment and contains operational, logistical, temporal, and market-specific attributes of shipping mode, order status, customer and order location, market region, and shipment characteristics. The data is based on real-world processes of digital supply chains and allows to model the operational risk in the large-scale and structured environment.

The preprocessing of data is done in line with strict leakage prevention and privacy protection measures. Before the modeling, the fields of personal identifiable information are eliminated to make the data meet the data protection requirements. Directly encoding post-outcome information, e.g. the actual length of delivery or profit measures affected by delay, are not used to prevent target leakage. Temporal characteristics are broken down into organized characteristics like order year, month, weekday and hour to represent seasonality and operational patterns without the loss of chronological integrity. Stratified sampling is used to divide the dataset into training and testing subsets to ensure the balance of classes in the results of late delivery. A 8020 split is used to baseline model and another time based split is used to train on the previous periods and test on the later periods. This time-based design aids in the assessment of predictive stability in the conditions of digital deployment.

3.3 Population and Sampling

The population of interest will include all the shipment transactions in the digital supply chain under analysis. Every unit of observation is a single order that is handled by the logistics system. The sampling method is non-probabilistic and archival because the research is based on a total operational dataset as opposed to survey-based or experimental data collection.

To ensure cross-validation efficiency, stratified subsampling has been used to produce computationally manageable subsets without any loss of proportional representation of late delivery outcomes. This is to ensure that the model training and validation is based on the underlying distribution of operational risk. Segment level performance appraisal also breaks down the data further to the market regions to determine whether predictive governance structures operate fairly in different global settings.

3.4 Data Analysis Techniques

The analysis incorporates predictive modeling, performance analysis, simulation-based governance analysis, and explainability. The supervised classification algorithms are applied in organised preprocessing pipelines to impute, scale and one-hot encode. Accuracy, precision, recall, F1-score and area under receiver operating characteristic curve (ROC-AUC) are used to measure model performance. Cross-validation gives the estimates of variance, whereas time-based ROC-AUC measures longitudinal robustness.

Analysis has been done on threshold optimization to analyze the impacts of probability cutoffs on the workload and cost structures of an organization. There are several levels of thresholds that are experimented to replicate other governance regimes. Each threshold is calculated in the number of flagged orders per day, false positives per shift, and true positive interventions. A cost model uses the cost of the intervention, the penalty cost of late delivery and the penalty cost of false review to determine the total cost of the organization under each threshold setting. This allows determining a cost-minimizing governance threshold and illustrates the way the algorithmic settings can be used as policy leverages.

In the segmented performance analysis, accuracy and recall of geographic markets are examined to determine equity in operations and fairness. Moreover, SHAP (SHapley Additive exPlanations) values are calculated in order to understand the contribution of features on the level of individual and groups. Directional SHAP analysis determines the effect of certain shipping modes or operational variables on delay risk, i.e. whether they increase or decrease delay risk. These understandings are converted into governance playbooks between drivers of risk and operational activities and accountable organizational entities. In such a combination of predictive and interpretive analytics, the research operationalizes the relationship between digitalization and institutional change.

3.5 Ethical Considerations

The study is based on responsible AI implementation and data management. All such data is anonymized before analysis and all modeling is done on anonymized transactional records. The leakage prevention protocols have been in place to make sure the predictive models are not based on the post-outcome variables that would artificially inflate the performance. Market fairness analysis deals with the issue of algorithmic bias and promotes fair digital governance.

4. Results

This part provides the empirical results of the AI-powered late-delivery risk model and shows how predictive digitalization is transforming the supply chain governance, the structure of operational workloads, and the institutional decision norms. The outcomes move through the technical validation stage to the governance calibration stage, workload redistribution stage, operational improvement stage, and explainability-based organizational transformation stage.

4.1 Predictive Performance and Model Validation

In order to develop a dependable base of AI-driven supply chain governance, three monitored classification frameworks were considered, including Logistic Regression, HistGradientBoosting, and XGBoost. Ensemble-based methods were always more effective than the linear baseline. HistGradientBoosting has the best discriminative performance with ROC-AUC of about 0.857 as opposed to 0.817 of Logistic Regression and XGBoost.

The comparison of the ROC curves of models is shown in Figure 1. The curve that corresponds to HistGradientBoosting always dominates the false-positive rates showing a better separation between late and on-time deliveries. This validates the appropriateness of ensemble-based learning to be used in operational implementation in digitally transformed supply chain systems.

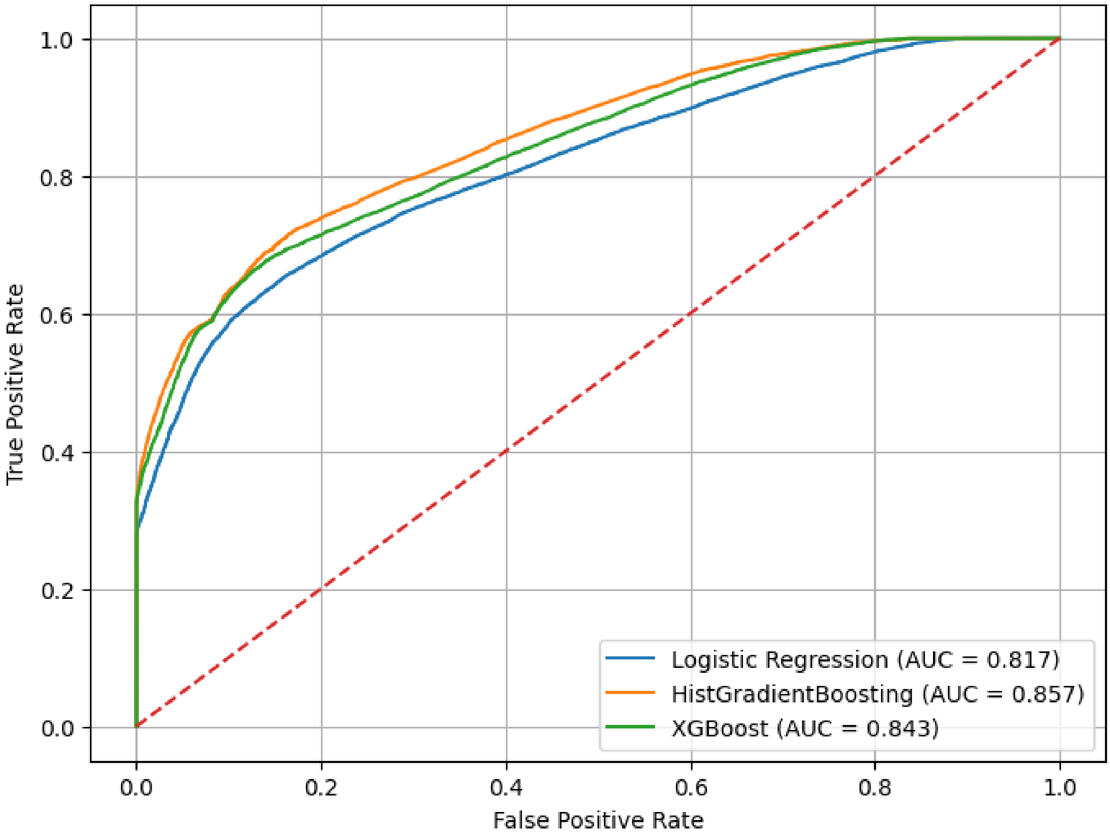

Figure 1. ROC Curve Comparison Across Predictive Models

The 5-fold cross-validation produced an average ROC-AUC of about 0.8146 with low variance, which proved that the model was stable and generalizable. In order to approximate real-world deployment conditions, a time split was used, where training was done on previous observations and testing was done on subsequent records. The time-based ROC-AUC of 0.8023 signifies low temporal drift but high sustained predictive accuracy, which proves that the AI system is not affected by changing conditions of its activity.

All these findings confirm the predictive core that is required of governance-level digital transformation.

4.2 Threshold Calibration as Governance Mechanism

Whereas predictive accuracy is a measure that determines technical reliability, digital governance is based on the translation of prediction scores into operational action. The classification threshold is a policy tool that identifies the intensity of intervention and allocation of resources.

The model yielded a large number of false negatives, which means that it missed high-risk shipments, at the traditional cutoff point of 0.50. The decrease in threshold minimizes risk that is missed but maximizes the workload of interventions.

Figure 2 shows the confusion matrix when the operational threshold is 0.40. False negatives are also much lower as compared to the default setting, which points to a change towards proactive risk mitigation. This, however, is at the expense of more false positives, which have to be reviewed further.

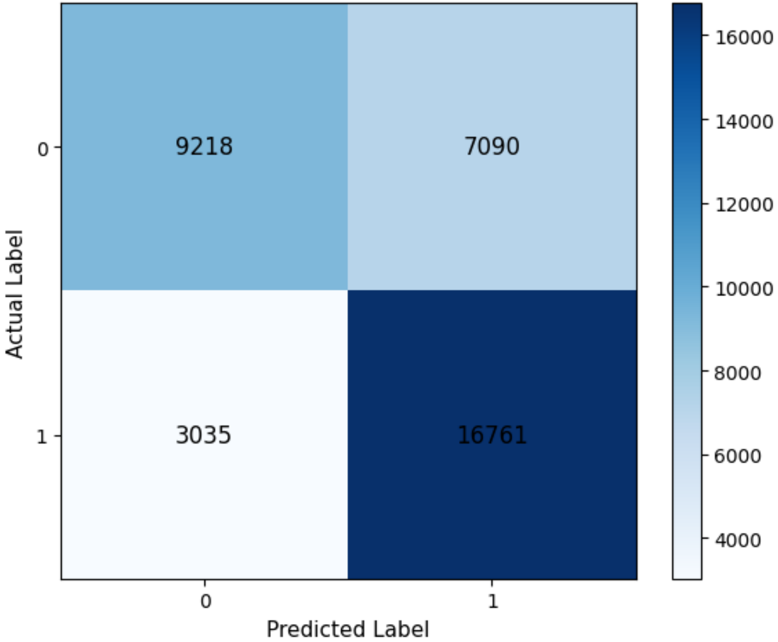

Figure 2. Confusion Matrix at Threshold = 0.40

To formalize this trade-off, a cost model incorporating late-delivery penalties, intervention expenses, and false-review costs was implemented.

Table 1. Cost–Benefit Analysis Across Probability Thresholds

Threshold	FN (Late Orders Missed)	Cost of Late Deliveries	Intervention Cost	False Review Cost	Total Estimated Cost
0.30	68	3,400	263,792	66,230	333,422
0.35	311	15,550	253,864	61,240	330,654
0.40	3,035	151,750	190,808	35,450	378,008
0.45	7,058	352,900	116,584	9,175	478,659
0.50	7,651	382,550	109,360	7,625	499,535
0.55	7,953	397,650	106,088	7,090	510,828

Table 1 shows that the choice of threshold is not only technical but is an example of a governance calibration mechanism regarding trade-offs between economic efficiency and reliability of the service. Through digitally transformed supply chains, threshold configuration realizes organizational risk tolerance.

4.3 Organizational Workload Redistribution

The AI-based governance transforms the workload patterns by shifting the operational focus on the high-risk shipments that are predicted. In order to evaluate organizational implications, flagged cases were converted into daily workload and shift-level workload.

Table 2. Organizational Workload Impact by Threshold

Threshold	Total Flagged Orders	Flagged per Day	False Positives	FP per Day	FP per Shift (3 shifts)
0.30	32,974	127.31	13,246	51.14	17.05
0.35	31,733	122.52	12,248	47.29	15.76
0.40	23,851	92.09	7,090	27.37	9.12
0.45	14,573	56.27	1,835	7.08	2.36

0.50	13,670	52.78	1,525	5.89	1.96
0.55	13,261	51.20	1,418	5.47	1.82

The cost-minimizing threshold (0.35) indicates that 31,733 shipments are flagged, which is 122.52 flagged cases per day. Based on three shifts per day, this would translate to an extra 15.76 false-positive reviews per shift.

Analysis of operation burden differences across thresholds show a significant difference. Thresholds that are lower raise the intensity of daily reviews, whereas higher thresholds decrease the work load but raise the risk that is missed. These results support the idea that AI implementation reorganizes the control practices and institutionalizes proactive monitoring processes in the supply chain operations. Digitalization thus changes the accuracy of predictions, as well as the organization of everyday operational attention.

4.4 Operational Improvement Through AI-Guided Intervention

To evaluate tangible performance gains, an intervention simulation was conducted for Standard Class shipments.

Table 3. Intervention Simulation Results for Standard Class

Metric	Value
Total Standard Class Orders	21,496
Historical Late Deliveries	8,156
High-Risk Orders Identified	10,860
Assumed Intervention Success Rate	50%
Estimated Late Deliveries Prevented	2,560
Percentage Reduction in Delays	31.39%

Out of 21,496 Standard Class orders, 8,156 had previously been late. The AI model found 10,860 dangerous Standard shipments. With a 50% rate of intervention success, about 2,560 late deliveries would be avoided, which is equivalent to a 31.39% decrease in Standard Class delays.

This finding shows that predictive digitalization is converted into quantifiable improvements in operational performance. The prioritization based on AI is, therefore, a tool of the institution to improve the reliability of the services and not to predict the outcomes.

4.5 Segmented Market Performance and Institutional Legitimacy

For AI governance to be sustainable, predictive performance must remain stable across geographic units.

Table 4. Predictive Performance Across Geographic Markets

Market	Orders	Precision	Recall
USCA	5,112	0.627	0.989
LATAM	10,365	0.608	0.985
Europe	9,966	0.612	0.980
Pacific Asia	8,328	0.616	0.986
Africa	2,333	0.613	0.984

Table 4 indicates that there are comparatively the same precision (mean 0.615) and high recall (mean 0.985) across markets. The small variance indicates that the predictive system does not affect certain areas unequally. This helps in the legitimization of the institutions and cross-regional acceptance of AI-based governance policies.

The consistency of cross-market performance supports the possibility of the standardization of digital decision-making frameworks in global supply chains.

4.6 Explainability and Structural Risk Drivers

To ensure interpretability and governance readiness, SHAP analysis was conducted to identify structural contributors to delay risk.

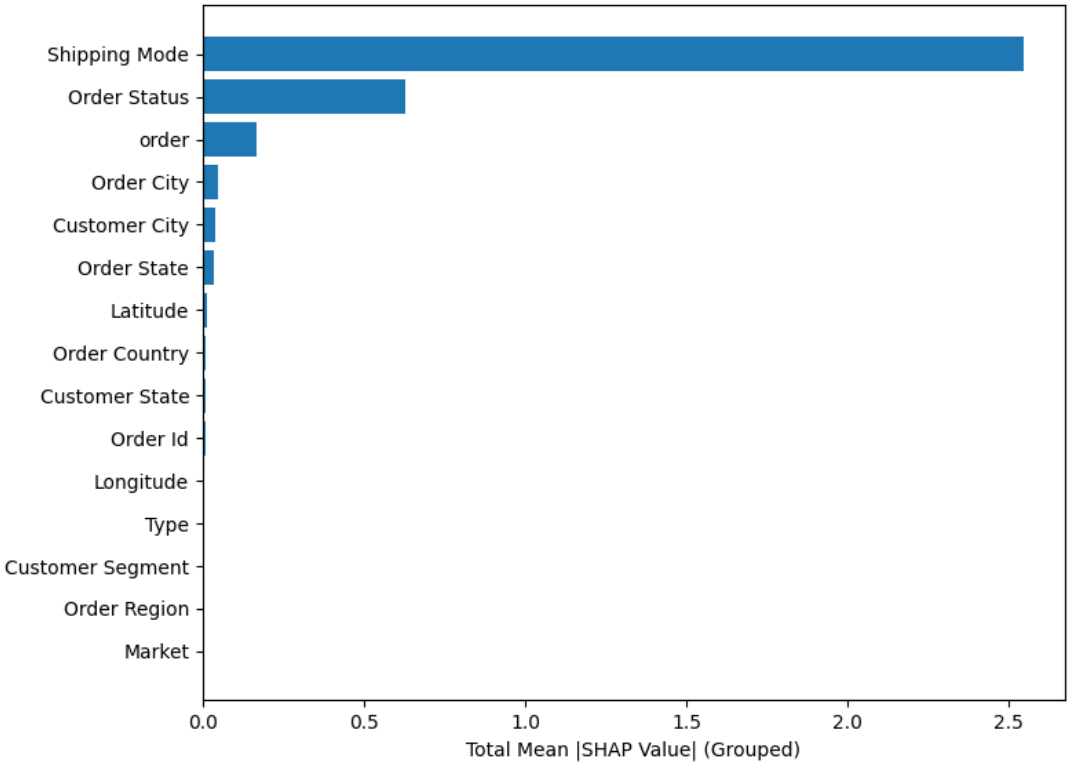

Figure 3. Grouped SHAP Feature Importance

Figure 3 indicates that the Shipping Mode is the most predictive factor of delay risk, then Order Status, and lastly, temporal demand factors. Geographic variables play a relatively smaller role on prediction outcomes. This implies that the delay risk is largely organized according to operational choices and workflow designs and not location-related factors.
In order to explain directionality, a narrow SHAP analysis of shipping modes was conducted.

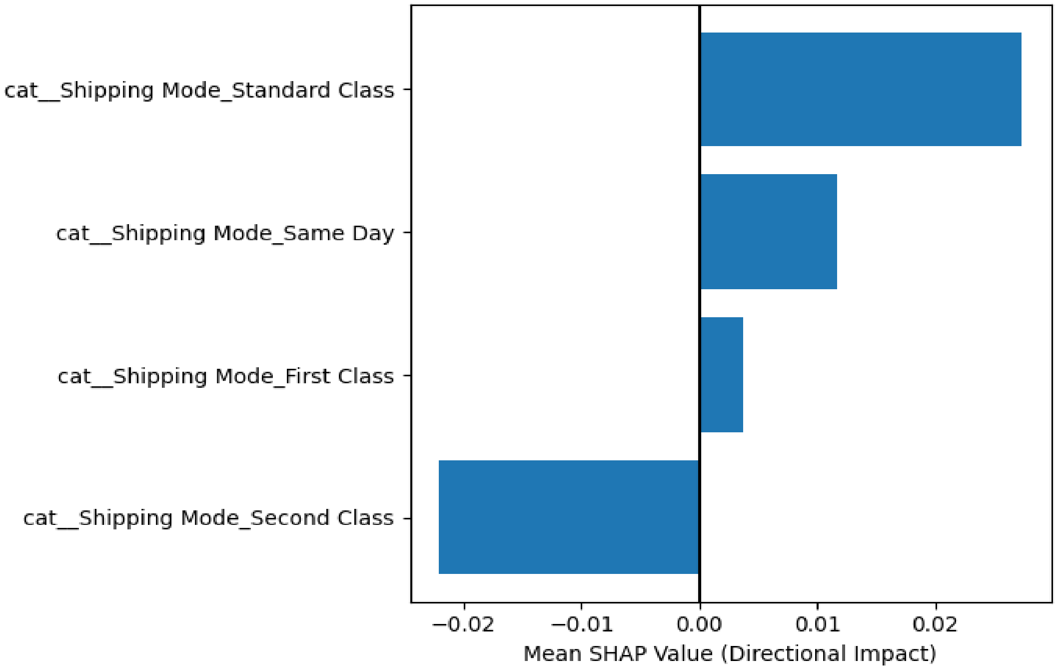

Figure 4. Directional SHAP Impact of Shipping Modes on Predicted Delay Risk.

The findings demonstrate that Standard Class and Same Day shipments raise the predicted delay probability, and First Class decreases the risk exposure. These results can offer practical governance suggestions to transport planning and prioritization strategies. Predictive analytics is converted into organized organizational knowledge through explainability.

4.7 Integrated Governance Dashboard

To synthesize digital transformation outcomes, a consolidated governance dashboard was developed.

Table 5. Integrated Governance Dashboard

Metric	**Value**
Optimal Threshold	0.35
Total Flagged Orders	31,733
Flagged per Day	122.52
False Positives per Shift	15.76
Estimated Total Cost	330,654
Time-Based ROC-AUC	0.8023
Mean Precision (Markets)	0.615
Mean Recall (Markets)	0.985

The dashboard summarises the most important deployment metrics, such as the cost-minimising threshold (0.35), flagged shipments (31,733), flagged per day (122.52), false positives per shift (15.76), estimated total cost (330,654), time-based ROC-AUC (0.8023), and segmented performance measures.

This is a combined perspective that proves that AI-related digitalization is not restricted to predictive optimization. It rebalances the trade-offs in governance, reorganizes workload distribution, improves the reliability of services, and institutionalizes the explainable decision-making in the supply chain operations.

5. Discussion

The results of this research prove that AI-based digitalization of supply chains is not confined to predictive accuracy, but it moves to the sphere of organizational governance and structural change. The predictive models had good discriminative ability, and the ensemble approaches had good ROC-AUC values and were able to perform well with time-based validation. Nonetheless, the more important input is the way these predictive outputs were converted into governance mechanisms. The discovery of a cost-reducing threshold shows that the operation of algorithmic calibration is a policy decision, which transforms the risk tolerance, intervention intensity, and operational attention. The findings of the modeling of workload per day and false positives per shift demonstrate how the implementation of AI can redistribute the capacity of the organization and institutionalize the routines of proactive oversight. This confirms the thesis that digitalization rearranges the norms of the functioning but does not necessarily lead to an increase in the accuracy of predictions.

The threshold analysis strengthens the cost sensitive theory of learning which focuses on comparing models with decision outcomes as opposed to overall accuracy (Elkan, 2001). The determination of an optimum threshold of 0.35 shows how predictive systems can be adjusted to fit the economic needs of organizations. In the same vein, the cost benefit model is consistent with the principles of decision-centric evaluation that emphasizes business-oriented model evaluation as opposed to purely statistical measures (Provost and Fawcett, 2013). In this regard, AI turns into a governance tool that equalizes economic efficiency and service reliability. Instead of using traditional default thresholds, organizations can also operationalize risk appetite explicitly by using data-driven calibration.

The explainability findings also highlight the structural change that is linked to AI-led operations. Analysis using SHAP showed that the most significant sources of delay risk were shipping mode and order status, indicating that performance variability is caused by systemic operational settings, and not entirely by geographic factors. This result echoes sociomaterial views in which technology and organizational practice are constitutive of each other (Orlikowski and Scott, 2008). The predictive system does not simply monitor the operational processes, but actively contributes to their restructuring by pointing out the structural bottlenecks and directing the approach to transport prioritization. The interpretability mechanisms transform the results of algorithms into governance playbooks, which supports the significance of transparency in the implementation of AI into the institutional decision-making process (Lundberg and Lee, 2017). Its organizational implications are on the structures of decision-making and the managerial positions. Since predictive thresholds are used to decide which shipment is given priority consideration, algorithmic scoring mediates managerial discretion to some extent. This substantiates claims that AI restructures power and organization in companies (Shrestha et al., 2019). The fact that the model allows decreasing the number of missed risky shipments by lower thresholds shows that proactive intervention can significantly enhance service reliability, as reflected by the simulated 31% decrease in the Standard Class delays. Meanwhile, more review work is an example of the automation-augmentation paradox, as AI not only improves performance but also creates new oversight demands (Raisch and Krakowski, 2021). In this way, AI-based digitalization helps to enhance the performance of operations and requires the structural adjustment of the workforce distribution and governance mechanisms.

The analysis of fairness among geographical markets adds to the responsible discussion of AI. The fact that the predictive system consistently recalls and has similar levels of precision across regions indicates that there is no disproportionate impact of the predictive system on any particular market, which is essential to the institutional legitimacy of global supply chains. Since there is growing interest in bias in machine learning systems, the incorporation of fairness diagnostics enhances the credibility of the governance of predictive deployment (Barocas et al., 2019). The interpretability, cost modeling, and segmentation analysis provide the AI system with the status of performance-enhancing and responsible.

As a manager, the findings indicate that AI implementation ought to be treated as a governance design and not a model development. The threshold setup has to be in line with the organizational capacity and economic priorities. The tools of explainability are needed to interpret the predictive results into working policies. Besides, the workload modeling proves that the digital transformation needs to be planned by systematically adjusting review routines and staffing strategies. Companies implementing AI-based business practices should thus invest not just in analytical systems, but also in a system of institutional alignment. In spite of these contributions, there are a number of limitations that need to be mentioned. To begin with, the research is based on archival transactional information and does not directly assess behavioral adaptation and employee perceptions of algorithmic control. Changes in culture are deduced based on the alterations in the structure and not empirically realized using the qualitative approach. Second, the intervention simulation presupposes a fixed rate of prevention, which can be different in real-life application. Third, the study is grounded in one supply chain dataset, which restricts its application to other industries or disruption situations.

Further studies should be conducted to expand this study by incorporating longitudinal field research studies to understand the interaction of employees with predictive systems with time. The alternative threshold policies can be tested under different capacity constraints using experimental designs. Moreover, the cross-industry comparative analysis would enhance the external validity and shed light on the sector-specific governance adaptation. Lastly, more profound implementation of fairness-based methods of learning might contribute to the equitable use in a variety of operational situations.

6. Conclusion

This paper has explored the role of AI-based predictive systems in digitalizing the supply chains by transforming the governance frameworks, operational processes, and decision-making standards. The results indicate that ensemble-based machine learning models have the capability of predicting the late-delivery risk with high levels of discriminative capabilities and temporal resilience. More to the point, the analysis indicates that probability threshold calibration is a governance mechanism, which impacts the cost exposure, workload allocation, and intervention intensity. The discovery of the cost-reducing point and the modelization of delay reduction represent the way predictive analytics can be converted into quantifiable operational changes. Additionally, the explainability analysis showed that delay risk is driven by structural operational factors, especially shipping mode and workflow status, which allows providing clear and policy-relevant governance information. This research has implications that are beyond technical optimization. The implementation of AI must be thought of as the redesign of an organization, which involves meticulous coordination of predictive thresholds, capacity limits and accountability frameworks. The managers are advised to make threshold selection a strategic policy choice and not a default technical choice. The further development of the understanding of AI-enabled supply chain transformation should be done through research on the longitudinal dynamics of adoption, how employees react to algorithmic supervision, and cross-industry testing of governance models.

References

1. Barocas, S., Hardt, M., & Narayanan, A. (2019). Fairness and machine learning.
2. Barrett, M., Davidson, E., Prabhu, J., & Vargo, S. L. (2015). Service innovation in the digital age: key contributions and future directions. *MIS quarterly*, *39*(1), 135-154.

3. Buyukozkan, G., & Gocer, F. (2018). Digital Supply Chain: Literature review and a proposed framework for future research. *Computers in industry*, *97*, 157-177.
4. Carbonneau, R., Laframboise, K., & Vahidov, R. (2008). Application of machine learning techniques for supply chain demand forecasting. *European journal of operational research*, *184*(3), 1140-1154.
5. Choi, T. M., Wallace, S. W., & Wang, Y. (2018). Big data analytics in operations management. *Production and operations management*, *27*(10), 1868-1883.
6. Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*.
7. Drummond, C., & Holte, R. C. (2006). Cost curves: An improved method for visualizing classifier performance. *Machine learning*, *65*(1), 95-130.
8. Elkan, C. (2001, August). The foundations of cost-sensitive learning. In *International joint conference on artificial intelligence* (Vol. 17, No. 1, pp. 973-978). Lawrence Erlbaum Associates Ltd.
9. Faraj, S., Pachidi, S., & Sayegh, K. (2018). Working and organizing in the age of the learning algorithm. *Information and organization*, *28*(1), 62-70.
10. Ivanov, D. (2021). Digital supply chain management and technology to enhance resilience by building and using end-to-end visibility during the COVID-19 pandemic. *IEEE Transactions on Engineering Management*, *71*, 10485-10495.
11. Ivanov, D., & Dolgui, A. (2020). Viability of intertwined supply networks: extending the supply chain resilience angles towards survivability. A position paper motivated by COVID-19 outbreak. *International journal of production research*, *58*(10), 2904-2915.
12. Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business horizons*, *61*(4), 577-586.
13. Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of management annals*, *14*(1), 366-410.
14. Leonardi, P. M. (2011). When flexible routines meet flexible technologies: Affordance, constraint, and the imbrication of human and material agencies1. *MIS quarterly*, *35*(1), 147-167.
15. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, *30*.
16. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM computing surveys (CSUR)*, *54*(6), 1-35.
17. Molnar, C. (2020). *Interpretable machine learning*. Lulu. com.
18. Orlikowski, W. J., & Scott, S. V. (2008). 10 sociomateriality: challenging the separation of technology, work and organization. *Academy of Management annals*, *2*(1), 433-474.
19. Provost, F., & Fawcett, T. Data Science for Business.
20. Queiroz, M. M., Pereira, S. C. F., Telles, R., & Machado, M. C. (2021). Industry 4.0 and digital supply chain capabilities: A framework for understanding digitalisation challenges and opportunities. *Benchmarking: an international journal*, *28*(5), 1761-1782.
21. Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of management review*, *46*(1), 192-210.
22. Sanders, N. R. (2016). How to use big data to drive your supply chain. *California management review*, *58*(3), 26-48.
23. Shashwatwork. (2019). *DataCo smart supply chain for big data analysis* [Data set]. Kaggle. Retrieved from https://www.kaggle.com/datasets/shashwatwork/dataco-smart-supply-chain-for-big-data-analysis
24. Shrestha, Y. R., Ben-Menahem, S. M., & Von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. *California management review*, *61*(4), 66-83.
25. Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of business research*, *70*, 356-365.
26. Yoo, Y., Henfridsson, O., & Lyytinen, K. (2010). Research commentary—the new organizing logic of digital innovation: an agenda for information systems research. *Information systems research*, *21*(4), 724-735.