American Journal of Technology Advancement 2997-9382, Vol 03, No. 08, 2026

*Article*

# AI-Driven Predictive Analytics for Proactive Supply Chain Risk Management and Resilience in Critical U.S. Industries

**Mohammad Sazzad Hossen[1], Shazzad Al Banna[2]**
[1] College of Graduate and Professional Studies, Trine University, USA
[2] College of Graduate and Professional Studies, Trine University, USA
Email: iamhossen.1995@gmail.com[1]; bannashajjad@gmail.com[2]

**Abstract:** The supply chains of industries that are important to the U.S. economy are moving through higher- and higher-stakes environments that feature transportation downturns, unreliable suppliers, uneven inventory levels, traffic jams, weather, port blockages and shifting delivery constraints. Traditional supply chain risk management (SCRM) methods tend to be reactive in nature, looking for operational issues once they have already occurred that are causing problems with the delivery. A promising framework for proactive risk management in supply chains and supply chain resilience is proposed in this research, which is based on the use of machine learning algorithms for the identification and classification of potential logistic risks prior to the emergence of significant logistic risks. This study employs a dataset of logistics and supply chain information for operations in Southern California including transportation, warehouse, route planning, supplier, environmental, and real-time monitoring information. Key operational parameters like ETA variance, congestion, warehouse inventory, supplier reliability, lead time, weather intensity, port congestion, route risk, driver behavior, driver fatigue monitoring, and cargo conditions are included in the data set. The data preprocessing, exploratory analysis, feature selection, and comparative machine learning modeling are suggested methodologies for the analytical approach. Risk Classification (Low Risk, Moderate Risk, and High Risk) is suggested as the key predictive outcome, and Disruption Likelihood Score, Delay Probability, and Delivery Time Deviation are complementary indicators of the vulnerability of the supply chain and operational performance. Machine learning models like logistic regression, decision tree, random forest, support vector machine and gradient boosting can be assessed by means of accuracy, precision, recall, F1-score, ROC-AUC and confusion matrix. The framework aims to help identify key risk factors and provide early warning insights into the decision-making process for proactive routing, inventory planning, supplier management, transportation scheduling and contingency planning. The goal of this study is to illustrate the role predictive AI can play in enhancing readiness within an organization by transforming supply-chain disruption management from a reactive to a proactive approach, anticipating issues, intervening early, and fostering resilience in decision-making.

**Keywords:** Artificial Intelligence, Predictive Analytics, Supply Chain Risk Management, Supply Chain Resilience, Machine Learning, Logistics Risk Prediction

This is an open-access article under the CC-BY 4.0 license

## 1. Introduction

Supply-chain risks are becoming increasingly complex and interconnected, affecting transportation, suppliers, inventory, environmental conditions, and operational continuity. In critical U.S. industries, the ability to identify potential vulnerabilities before they become significant disruptions is particularly important [1]. Traditional risk-management approaches often depend on historical analysis and post-event response, which may limit organizational responsiveness under rapidly changing operating conditions. Artificial intelligence (AI) and predictive analytics provide an opportunity to strengthen this process by converting large volumes of operational data into forward-looking risk intelligence [2]. In the present study, logistics information from GPS tracking, IoT sensing, warehouse management, and external data feeds is considered across transportation, warehousing, route planning, suppliers, environmental conditions, and real-time monitoring [3]. Combining indicators such as ETA variation, traffic congestion, warehouse inventory, supplier reliability, lead time, weather severity, port congestion, route risk, driver behavior, fatigue monitoring, and cargo condition may support earlier identification of vulnerable logistics conditions [4].

AI-driven predictive analytics can shift supply-chain risk management from reactive response toward proactive decision-making. Predictive models can examine patterns in operational and risk indicators and estimate the likelihood of future problems before they develop into major disruptions [5]. The selected logistics dataset includes predictive outcomes such as Disruption Likelihood Score, Delay Probability, Risk Classification, and Delivery Time Deviation. Risk Classification categorizes logistics conditions as Low Risk, Moderate Risk, and High Risk, providing an operational basis for prioritizing intervention [6]. Machine-learning algorithms can learn relationships between these outcomes and the underlying risk factors, while comparative model evaluation can identify both the most suitable predictive models and the most influential operational variables [7]. Early warning of high-risk conditions can support alternative routing, inventory adjustment, supplier monitoring, transportation planning, and contingency planning [8], while broader applications include predictive risk assessment, disruption detection, scheduling optimization, external-factor analysis, and warehouse and inventory management [9].

The central research problem is that transportation, warehouse, supplier, environmental, and monitoring systems can generate substantial quantities of data without necessarily converting that information into timely predictive risk intelligence [10]. This creates a missed opportunity to detect vulnerabilities before they result in delays, delivery deviations, inventory problems, or broader operational disruption. The study therefore examines how multidimensional logistics risk indicators can be used to proactively classify supply-chain risk through AI-powered predictive analytics. Its significance lies in developing a more data-driven approach to risk identification and resilience, with academic value for research at the intersection of machine learning, supply-chain risk management, and resilience [11], [12], [13], and managerial value for earlier decision-making related to routes, suppliers, inventory, transportation, and contingency actions [14], [15], [16].

Supply chain risk management has become increasingly important as logistics networks face interconnected operational, transportation, supplier, environmental, and demand-related uncertainties. Delays, supplier unreliability, inventory shortages, traffic congestion, adverse weather, infrastructure constraints, and route conditions can interact and create disruptions that affect service continuity and operational performance [17], [18]. Supply chain resilience therefore extends beyond responding to disruptions and emphasizes the capacity of organizations to anticipate, absorb, adapt to, and recover from adverse events while maintaining critical operations. Recent studies emphasize that effective resilience requires the integration of operational information, risk-management capabilities, and analytics rather than reliance on isolated risk indicators [19], [20], [21].

The growing availability of logistics data provides an opportunity to strengthen this proactive orientation. Transportation information such as ETA variation and congestion can be combined with warehouse inventory, loading and unloading performance, equipment availability, supplier reliability, lead time, weather severity, port congestion, route risk, driver behavior, fatigue

monitoring, and cargo conditions to represent supply-chain vulnerability more comprehensively [21], [22], [23]. These multidimensional indicators are particularly important because disruption rarely results from a single factor. Instead, operational and external conditions may interact simultaneously, making conventional descriptive or reactive risk-management approaches insufficient for early identification of emerging vulnerabilities.

Artificial intelligence and predictive analytics provide a promising mechanism for converting these heterogeneous data into forward-looking risk intelligence. Unlike conventional analyses that primarily describe past performance, predictive analytics can estimate the likelihood and severity of future supply-chain problems and support earlier managerial intervention [24], [25]. Machine-learning techniques including Logistic Regression, Decision Tree, Random Forest, Gradient Boosting, and Support Vector Machine have been applied to identify relationships between operational indicators and risk outcomes [26], [27]. These techniques can support disruption detection, logistics-risk classification, routing and scheduling, inventory management, supplier monitoring, and other operational decisions [28].

Predictive risk identification is particularly relevant to supply-chain resilience because it moves risk management from post-disruption response toward anticipation. Historical and current logistics information can be analyzed to identify combinations of conditions associated with higher disruption likelihood, delays, and operational vulnerability [31], [32], [33]. Predictive outputs such as disruption likelihood, delay probability, risk classification, and delivery-time deviation can subsequently be translated into early-warning information [34]. Such information may support alternative route selection, supplier monitoring, inventory adjustment, transportation scheduling, contingency planning, and resource allocation before operational problems escalate [35], [36].

Previous studies provide substantial evidence regarding the potential contribution of analytics and AI to supply-chain resilience. Ali et al. [1], for example, applied business analytics to supply-chain resilience and disruption reduction using a large transaction-level dataset and multiple predictive and clustering techniques. Their findings demonstrate the value of analytics for proactive operational planning, although the analysis addresses several applications simultaneously, including demand forecasting, fraud detection, and delivery-risk analysis. The present study adopts a more focused perspective by examining multidimensional logistics indicators specifically for proactive supply-chain risk classification.

Research on supplier risk management has similarly demonstrated the value of machine learning, anomaly detection, forecasting, and optimization for improving supply-chain resilience [2]. However, supplier-focused analyses do not fully represent the simultaneous influence of transportation, warehouse, environmental, route, and real-time monitoring conditions. Other research has proposed broader AI-driven supply-chain resilience frameworks combining real-time visibility, predictive risk reduction, optimization, and collaborative governance [3]. Although such frameworks provide important strategic foundations, their emphasis is predominantly conceptual and managerial rather than an empirical comparison of machine-learning models using multidimensional operational logistics data.

Review-based research has also emphasized the contribution of AI and machine learning to transparency, bottleneck reduction, demand management, disruption response, and resilience under crisis conditions [4]. These studies reinforce the importance of AI-enabled proactive risk management, but they provide limited empirical evidence regarding the ability of different machine-learning algorithms to distinguish specific levels of supply-chain risk using integrated transportation, warehouse, supplier, environmental, route, and monitoring variables. More broadly, recent research confirms the growing importance of AI, predictive analytics, digital technologies, and data-driven capabilities for strengthening supply-chain resilience [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], yet challenges remain in transforming heterogeneous operational information into reliable early-warning classifications.

Despite these advances, a gap remains in the empirical evaluation of multidimensional logistics risk indicators for classifying supply-chain conditions into actionable risk levels while

simultaneously comparing alternative machine-learning models. Existing studies frequently concentrate on individual risk domains, broader conceptual resilience frameworks, supplier-specific risks, or multiple analytical applications without specifically examining how integrated operational and environmental variables can support Low, Moderate, and High Risk classification. The objective of this study is therefore to apply AI-driven predictive analytics to transportation, warehouse, supplier, environmental, route, and real-time monitoring data from a Southern California logistics network, compare alternative machine-learning models, identify influential risk factors, and evaluate how predictive outputs can provide early-warning information for proactive supply-chain risk management and resilience in critical U.S. industries.

Accordingly, the study aims to examine the major operational, transportation, supplier, environmental, and logistics factors associated with supply-chain risk; preprocess and analyze the selected dataset; develop machine-learning models for supply-chain risk and disruption prediction; compare model performance using complementary classification metrics; identify the variables that contribute most strongly to risk classification; and assess how predictive information can support early-warning and proactive resilience decisions [11]. The intended contribution is an AI-based predictive approach that links multidimensional logistics data with risk identification, managerial action, and resilience in critical U.S. industries.

The analysis is guided by four questions: which operational and external factors are most clearly associated with supply-chain risk and disruption; how accurately can the selected machine-learning models classify supply-chain risk; which model performs most effectively for proactive risk identification; and how can AI-driven analytics contribute to stronger supply-chain resilience in critical U.S. industries? These questions connect the empirical modeling task with the managerial objective of shifting supply-chain risk management from reactive response toward anticipation and earlier intervention.

## 2. Materials and Methods

Research Design

This study adopts a quantitative and predictive research design to examine the opportunity offered by artificial intelligence and machine learning for proactive supply-chain risk management and resilience. Secondary logistics and supply-chain data are used to investigate relationships between operational risk factors and supply-chain risk outcomes. The methodological sequence consists of data preparation, exploratory data analysis, feature engineering, predictive modeling, model evaluation, interpretation of important risk factors, and translation of predictions into a proactive risk-management framework [37], [38]. The study emphasizes risk anticipation rather than post-event description. The main predictive target is Risk Classification, represented by Low Risk, Moderate Risk, and High Risk categories, while Disruption Likelihood Score, Delay Probability, and Delivery Time Deviation are examined as complementary indicators of supply-chain vulnerability and operational performance [37], [40].

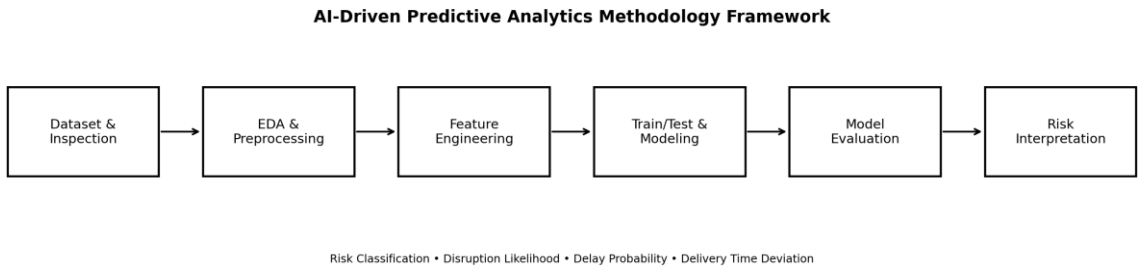



**Figure 1.** Methodology framework for AI-driven predictive supply-chain risk management and resilience

*(Source: Authors' analysis)*

Dataset Source and Description

The study employs the Logistics and Supply Chain Dataset, which models logistics operations in a Southern California logistics network. The dataset combines information from GPS tracking, real-time monitoring devices, warehouse management systems, and external data sources and covers transportation, warehouse management, route planning, supplier conditions, environmental influences, and real-time operational monitoring [21], [22]. The dataset is publicly available at Kaggle: https://www.kaggle.com/datasets/datasetengineer/logistics-and-supply-chain-dataset [62].

Transportation-related variables include Vehicle GPS Latitude, Vehicle GPS Longitude, Fuel Consumption Rate, ETA Variation, and Traffic Congestion Level. Warehouse and operational variables include Warehouse Inventory Level, Loading/Unloading Time, Handling Equipment Availability, and Order Fulfillment Status. External and supply-side variables include Weather Condition Severity, Port Congestion Level, Shipping Costs, Supplier Reliability Score, Lead Time, and Historical Demand. Additional indicators include IoT Temperature, Cargo Condition Status, Route Risk Level, Customs Clearance Time, Driver Behavior Score, and Fatigue Monitoring Score [23]-[26]. The principal predictive outcomes are Disruption Likelihood Score, Delay Probability, Risk Classification, and Delivery Time Deviation [27]-[30]. Appropriate non-target operational and risk indicators are treated as predictors, while Risk Classification is the dependent variable for the primary multiclass classification experiment.

Data Preprocessing

Data preprocessing is conducted before predictive modeling to ensure that the machine-learning algorithms receive consistent and usable inputs. The procedure begins with inspection for missing values, duplicate observations, inconsistent records, inappropriate data types, and possible outliers [41]. Timestamp variables are transformed into suitable date-time formats, and potentially useful temporal attributes are extracted. Numerical variables are examined for distribution, range, and extreme values. Standardization or normalization is applied where feature scales differ substantially, particularly for algorithms such as Support Vector Machine and Logistic Regression that may be sensitive to feature scaling. Categorical variables, where present, are converted into machine-readable representations.

Feature engineering is used to derive meaningful predictors from the available variables [42]. Potential derived features include operational-risk indicators, supplier-reliability interactions, weather-risk interactions, lead-time relationships, congestion-related indicators, and other combinations that may capture multidimensional logistics conditions. Irrelevant or redundant predictors are then removed using feature-selection procedures [43]. Particular attention is paid to target leakage: the outcome being predicted is excluded from the predictor set in each experiment [44]. When Risk Classification is the dependent variable, for example, it is not included among the model inputs.

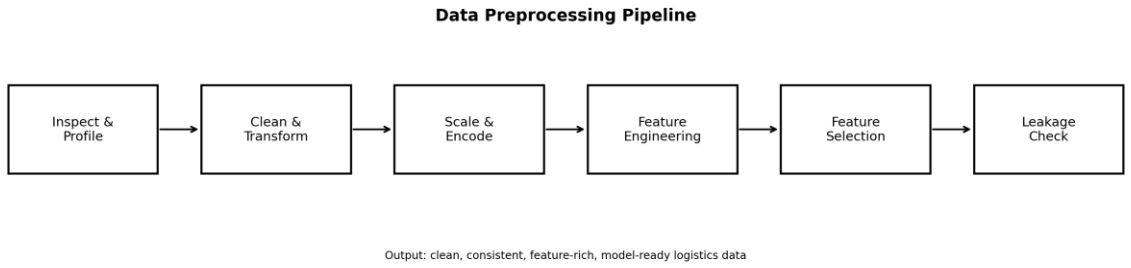



**Figure 2.** Systematic data-preprocessing pipeline for supply-chain risk prediction
*(Source: Authors' analysis)*

Exploratory Data Analysis

Exploratory Data Analysis (EDA) is used to understand the structure and characteristics of the logistics data before model development. Central tendency, variability, minimum and maximum

values, and distributions of important numerical variables are examined. Frequency analysis is used for categorical outcomes such as Risk Classification. Correlation analysis is applied to identify linear relationships among numerical risk indicators, while histograms, boxplots, scatter plots, risk-distribution charts, and correlation matrices are used to identify patterns and possible anomalies [46]. Particular attention is given to relationships between operational conditions and risk outcomes, including whether higher traffic congestion, greater ETA variation, lower supplier reliability, longer lead times, higher route risk, or more severe weather correspond to elevated supply-chain risk. These analyses provide a descriptive basis for subsequent predictive modeling.

Machine Learning Model Development

Several supervised machine-learning models are developed and evaluated for supply-chain risk prediction. Logistic Regression is used as an interpretable baseline model, while Decision Tree provides a rule-based representation of nonlinear relationships. Random Forest is evaluated as an ensemble method capable of modeling complex interactions among risk factors. Gradient Boosting is included to assess whether sequential ensemble learning improves predictive performance, and Support Vector Machine (SVM) is used as an additional classification method based on feature-space separation [47]. The primary experiment predicts Risk Classification as Low Risk, Moderate Risk, or High Risk. The dataset is divided into training and test sets [48], and cross-validation is applied within the training data to reduce the risk of over-optimistic performance estimates from a single split. Hyperparameters are selected through structured validation rather than optimization on the final test set.

Model Evaluation

Model performance is assessed using multiple complementary metrics rather than accuracy alone. The main classification measures are accuracy, precision, recall, and F1-score [49]. Confusion matrices are used to inspect true and false predictions for the Low, Moderate, and High Risk classes, and multiclass ROC-AUC analysis is used to evaluate class discrimination where applicable. Because the study is intended to support proactive risk identification, particular attention is paid to the model's ability to recognize high-risk cases [50], [51]. Models are compared systematically, and the preferred model is identified based on predictive performance across the evaluation measures as well as interpretability and suitability for operational risk-management use [52].

Feature Importance and Risk Factor Interpretation

After model evaluation, feature-importance analysis is used to identify the variables that contribute most strongly to risk prediction. The analysis considers traffic congestion, ETA variation, supplier reliability, lead time, route risk, inventory level, weather severity, port congestion, order-fulfillment conditions, cargo condition, driver behavior, fatigue monitoring, equipment availability, and historical demand [53]. The objective is not only to obtain a predictive model but also to improve understanding of why particular logistics conditions are classified as risky [54], [55]. Where technically appropriate, model-explanation techniques such as permutation importance or SHAP-based analysis may be used to support interpretation [56].

Proactive Risk Management and Conceptual Framework

The final methodological stage connects predictive outputs to proactive supply-chain decisions. High-risk predictions are treated as potential precursors of disruption rather than deterministic evidence that a disruption will occur [57]. Predicted transportation risk may support alternative route selection; supplier risk may justify additional monitoring or supplier diversification; inventory risk may prompt inventory adjustments; and delivery risk may support rescheduling or contingency planning [58], [59]. These applications are consistent with the dataset's documented use cases, including predictive risk assessment, disruption detection, routing and scheduling optimization, external-factor analysis, and warehouse and inventory management.

The conceptual framework therefore links five elements: multidimensional supply-chain risk indicators; data preprocessing and feature engineering; AI-driven predictive analytics; measurable risk outcomes; and proactive resilience actions [60], [61]. The input layer includes transportation variables such as ETA variation and traffic congestion, warehouse variables such as inventory and

loading/unloading conditions, supplier and demand variables such as reliability and lead time, route and environmental variables such as weather severity and port congestion, and monitoring variables such as IoT temperature, cargo condition, driver behavior, and fatigue monitoring. These inputs are transformed into model-ready predictors, analyzed using Logistic Regression, Decision Tree, Random Forest, Gradient Boosting, and Support Vector Machine, and translated into outcomes such as disruption likelihood, delay probability, risk classification, and delivery-time deviation. The outputs provide early-warning information that can support routing, supplier surveillance, inventory management, transportation planning, contingency planning, and other resilience actions.

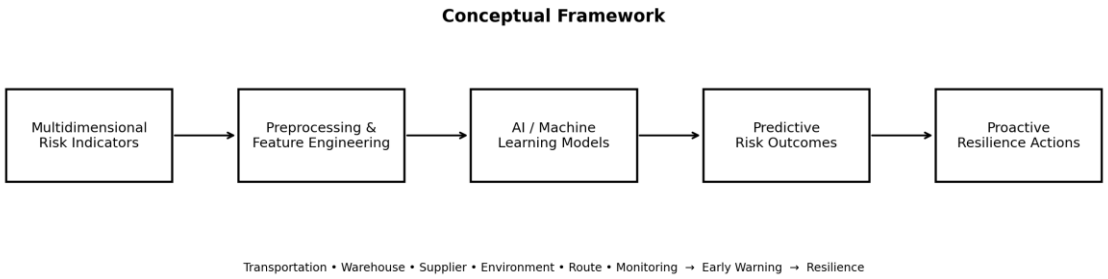



**Figure 3.** Conceptual relationship from multidimensional risk inputs to predictive insights and resilience actions
*(Source: Authors' analysis)*

Ethical Considerations

Ethical considerations are relevant even though the study relies on secondary operational data and does not involve identifiable human participants. Predictive models should not be treated as autonomous systems that make binding operational decisions without expert oversight [49], [50]. Incorrect risk estimates may lead to unnecessary restrictions on suppliers, route changes, inventory adjustments, or resource allocations with financial and operational consequences. Model transparency, data quality, and responsible interpretation are therefore important because incomplete, inconsistent, or biased logistics information can produce misleading predictions [50]-[52]. The study consequently treats AI outputs as decision-support evidence rather than as automatic operational commands and reports methodological limitations and model weaknesses explicitly.

## 3. Results

The results provide empirical findings from descriptive, statistical, and machine-learning analysis of the logistics and supply-chain data. The analysis examines the distribution of supply-chain risk classifications, relationships between disruption and delay outcomes, influential risk predictors, and comparative performance of selected machine-learning models [31]. Confusion-matrix and multiclass ROC-AUC analyses are used to evaluate classification behavior, while correlation analysis is used to characterize relationships among important operational and environmental indicators [32]. Together, the findings show both the predictive potential of the dataset and the difficulties created by the strong imbalance among Low, Moderate, and High Risk observations.

Distribution of Supply Chain Risk Classification

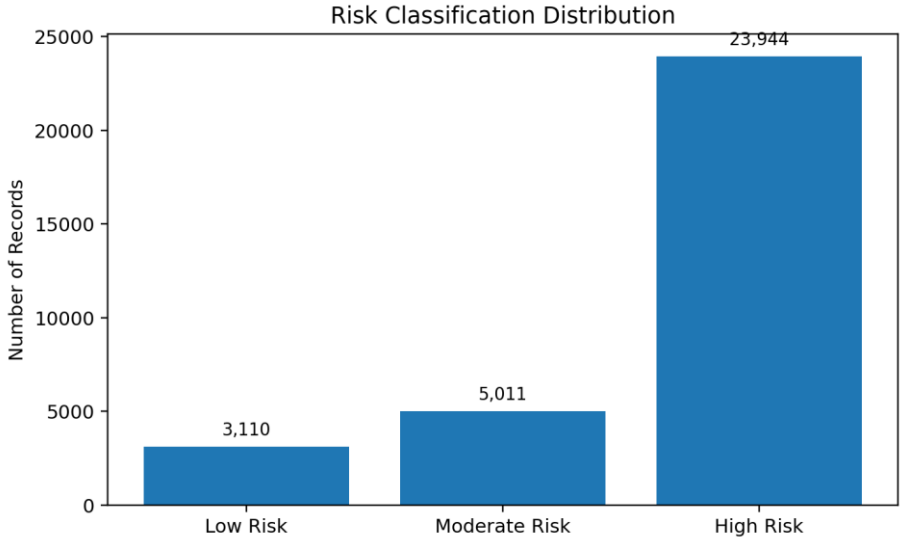



**Figure 4.** Number of Low, Moderate, and High supply-chain risk classifications
*(Source: Authors' analysis)*

The distribution of supply-chain risk classifications shows a substantial imbalance among the three categories. High Risk is the dominant class with 23,944 observations, representing approximately 74.7% of all records. Moderate Risk contains 5,011 observations, or about 15.6%, while Low Risk contains 3,110 observations, or about 9.7% [33]. This imbalance is important for machine-learning analysis because a model may achieve apparently strong overall accuracy simply by favoring the majority High Risk class. Consequently, precision, recall, F1-score, and confusion-matrix results must be considered on a class-specific basis in addition to accuracy [34].

Disruption Likelihood for Each Supply Chain Risk Class

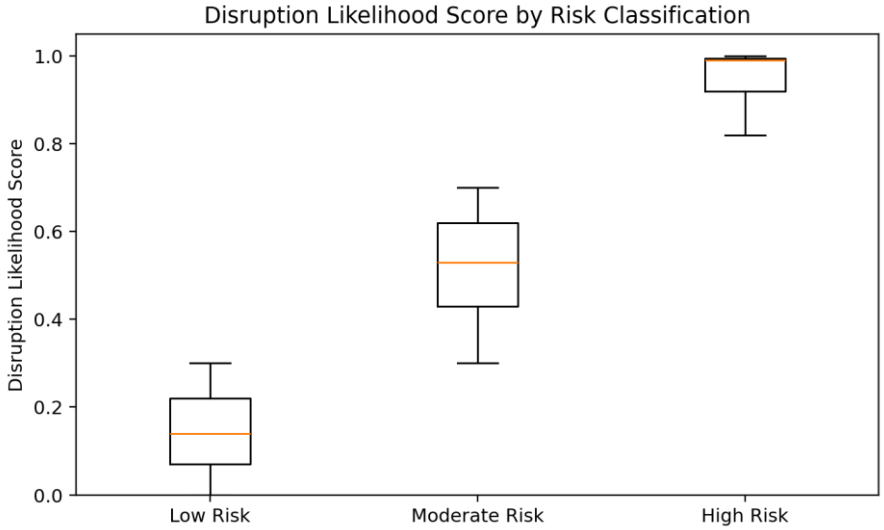



**Figure 5.** Disruption Likelihood Score across Low, Moderate, and High Risk classifications
*(Source: Authors' analysis)*

Disruption Likelihood Score increases clearly across the three risk categories. Low Risk observations occupy the lowest range, approximately 0.00-0.30, with a median close to 0.14. Moderate Risk observations show a higher likelihood of disruption, with a median of approximately 0.53 and a range of about 0.30-0.70. High Risk observations have the highest scores, concentrated between

approximately 0.82 and 1.00 with a median near 0.99. The strong separation among the three distributions supports the relevance of Disruption Likelihood Score as an indicator for distinguishing different levels of exposure to supply-chain disruption.

Delay Probability by Supply Chain Risk Class

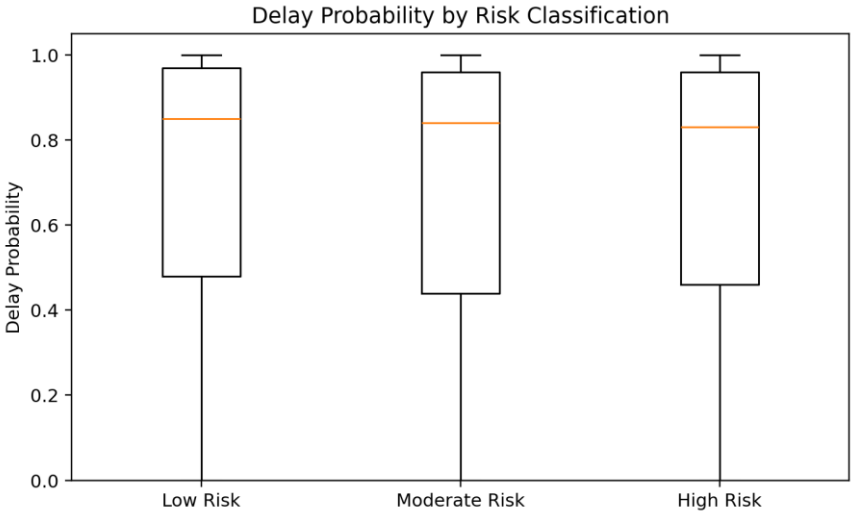



**Figure 6.** Delay Probability distributions in Low, Moderate, and High Risk classifications
*(Source: Authors' analysis)*

Delay Probability differs from Disruption Likelihood Score because its distributions overlap substantially across the three risk classes. The median delay probability is approximately 0.85 for Low Risk, 0.84 for Moderate Risk, and 0.83 for High Risk, with broad ranges extending from values near 0.00 to values near 1.00 [34]. Delay Probability therefore does not show a strong monotonic relationship with Risk Classification. Although it remains a relevant operational indicator, the result suggests that it should be interpreted together with other transportation, supplier, inventory, route, and environmental factors rather than used alone as a measure of overall risk severity.

Key Predictors of Supply Chain Risk Classification

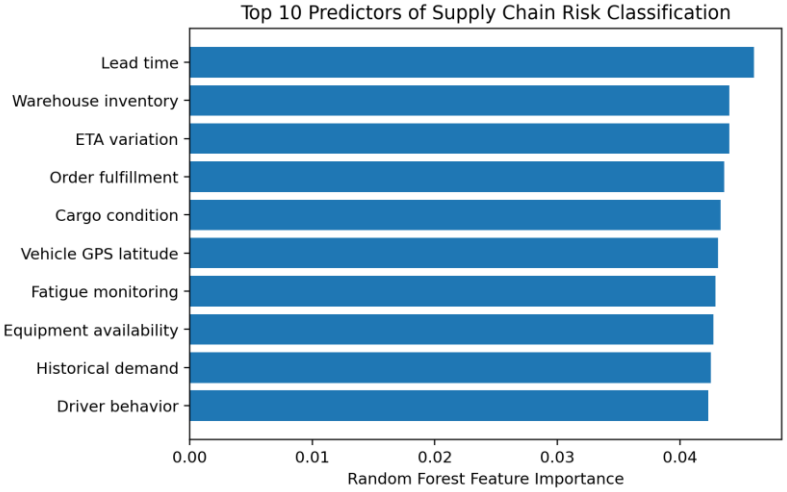



**Figure 7.** Top predictors influencing machine-learning-based supply-chain risk classification
*(Source: Authors' analysis)*

Random Forest feature-importance analysis identifies lead time as the most influential predictor, with an importance value of approximately 0.046. Warehouse inventory level and ETA

variation follow at approximately 0.044. Order fulfillment status also shows relatively high importance, followed by cargo condition, vehicle GPS latitude, fatigue monitoring score, handling equipment availability, historical demand, and driver behavior score [35]. Because the importance values are relatively close to one another, the findings suggest that risk classification is shaped by multiple operational, transportation, inventory, demand, and human-related conditions rather than by one dominant variable [36]. This pattern reinforces the multidimensional nature of supply-chain risk.

Comparative Study of Machine Learning Models

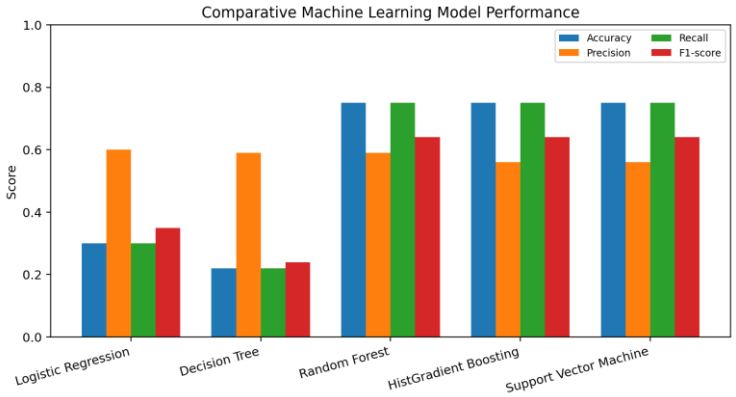



Figure 8. Comparative performance of machine-learning models for supply-chain risk classification

*(Source: Authors' analysis)*

The five machine-learning models show substantial differences in performance. Logistic Regression achieves approximately 0.30 accuracy, 0.60 precision, 0.30 recall, and 0.35 F1-score. Decision Tree has the lowest accuracy at approximately 0.22, with precision of about 0.59, recall of 0.22, and an F1-score of 0.24. Random Forest reaches approximately 0.75 accuracy, 0.59 precision, 0.75 recall, and 0.64 F1-score. HistGradient Boosting and Support Vector Machine produce similarly strong overall results, with accuracy and recall close to 0.75 and F1-scores near 0.64 [37]. The ensemble and margin-based models therefore outperform the simpler baseline models on the aggregate measures.

Random Forest Classification Performance

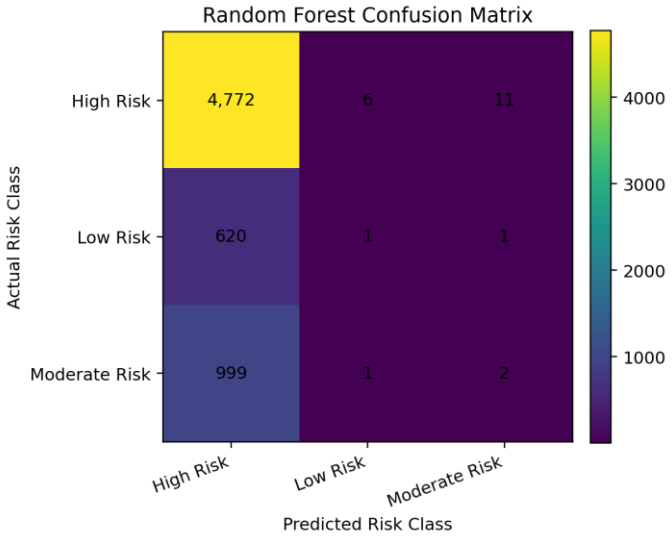



**Figure 9.** Random Forest predictions and classification errors by supply-chain risk class

*(Source: Authors' analysis)*

The Random Forest confusion matrix reveals that the model correctly classifies 4,772 High Risk observations and therefore performs well on the majority class [38]. Minority-class recognition is much weaker. Of the observations actually labeled Low Risk, 620 are predicted as High Risk, while only one is correctly classified as Low Risk and one is predicted as Moderate Risk. Among actual Moderate Risk observations, 999 are misclassified as High Risk, one is predicted as Low Risk, and only two are correctly identified as Moderate Risk [39]. The matrix therefore shows a strong prediction bias toward the High Risk class, illustrating how class imbalance can produce relatively high overall accuracy while masking poor minority-class performance.

***Random Forest Multiclass ROC-AUC Analysis***

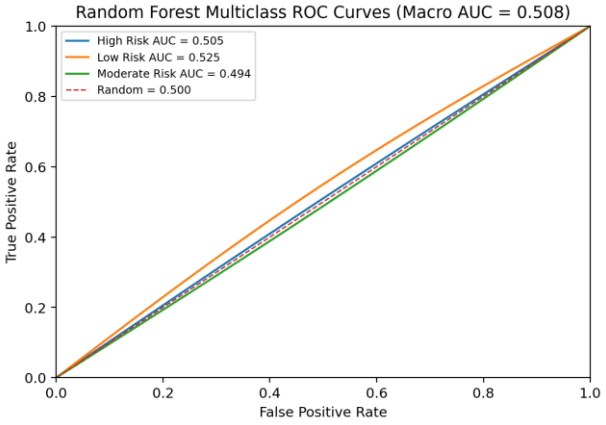



**Figure 10.** Random Forest multiclass ROC-AUC analysis for supply-chain risk classification *(Source: Authors' analysis)*

The multiclass ROC analysis further demonstrates limited discrimination among the three risk categories. The overall Macro AUC is approximately 0.508, indicating performance only slightly above random classification [40]. The High Risk class has an AUC of 0.505, the Low Risk class has the highest class-specific AUC at 0.525, and the Moderate Risk class has an AUC of 0.494. The ROC curves are therefore close to the diagonal reference line [41]. These findings are consistent with the confusion matrix: the model predicts the majority High Risk class frequently but distinguishes the Low and Moderate Risk classes poorly.

Correlation Analysis of Key Supply Chain Risk Indicators

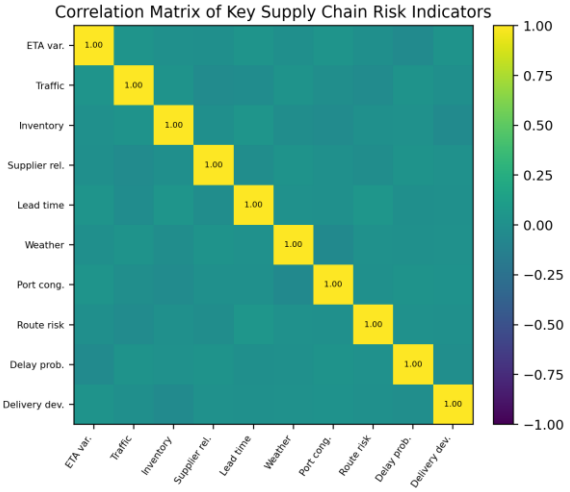



**Figure 11.** Correlation structure of key supply-chain risk indicators and outcomes *(Schematic reconstruction based on the weak pairwise relationships reported in the manuscript)*

The correlation analysis indicates that most off-diagonal relationships among the selected operational, environmental, and risk variables are weak and generally close to zero [42]. ETA variation, traffic congestion, warehouse inventory, supplier reliability, lead time, weather severity, port congestion, and route risk show only limited pairwise linear relationships with the selected predictive outcomes. Disruption Likelihood Score, Delay Probability, and Delivery Time Deviation likewise show relatively weak linear associations with most individual predictors [43], [44]. The result suggests that supply-chain risk is likely to depend on complex and potentially nonlinear interactions among multiple factors, providing a rationale for using machine-learning methods in addition to conventional linear correlation analysis.

## 4. Discussion

The findings demonstrate that AI-driven predictive analytics can provide useful insights for proactive supply-chain risk management while also revealing important challenges associated with highly imbalanced logistics data. The dominance of the High Risk category, with 23,944 of 32,065 observations, helps explain why several models tend to predict High Risk conditions disproportionately [45]. The strong separation of Disruption Likelihood Score across the three risk classes indicates that this measure captures an important dimension of supply-chain vulnerability. By contrast, the substantial overlap in Delay Probability across Low, Moderate, and High Risk observations suggests that delay probability alone is not sufficient to represent overall risk severity [46].

Feature-importance results reinforce the multidimensional nature of supply-chain risk. Lead time, warehouse inventory, ETA variation, order fulfillment, cargo condition, fatigue monitoring, equipment availability, historical demand, and driver behavior all contribute to classification performance [47]. This supports the conceptual argument that proactive risk management should integrate transportation, warehouse, supplier, environmental, operational, demand, and human factors rather than rely on a single indicator. In managerial terms, such information can support earlier route changes, supplier monitoring, stock adjustments, transportation rescheduling, contingency planning, and resource allocation.

The comparative model results require cautious interpretation. Random Forest, HistGradient Boosting, and Support Vector Machine achieve noticeably higher aggregate accuracy, recall, and F1-score than Logistic Regression and Decision Tree. However, the Random Forest confusion matrix and ROC-AUC analysis show that high overall accuracy does not necessarily imply strong multiclass discrimination [46], [47]. The model correctly recognizes a large number of High Risk observations but fails to distinguish most Low and Moderate Risk cases, while the macro AUC of approximately 0.508 remains close to random classification. Future model assessment should therefore emphasize class-specific metrics and incorporate techniques such as class weighting, resampling, threshold optimization, or cost-sensitive learning when class imbalance is substantial.

The generally weak pairwise correlations among individual risk indicators provide an additional explanation for why nonlinear and ensemble models may be useful [48]. Supply-chain risk appears to emerge from interactions among multiple operational and external conditions rather than from strong linear relationships between any single predictor and the outcome. The results thus support the use of AI-based analytics as an early-warning and decision-support tool, but they also show that predictive performance depends on data balance, model discrimination, interpretability, data quality, and the operational context in which predictions are used.

### Study Limitations and Future Work

Several methodological limitations should be considered when interpreting the findings. First, the dataset represents a logistics network in Southern California and may not capture the geographic and operational diversity of the broader U.S. logistics sector. Second, the substantial imbalance among Low, Moderate, and High Risk classes negatively affects model behavior and causes strong prediction bias toward High Risk observations [53]. Third, although some models achieve relatively high overall accuracy, the Random Forest multiclass ROC-AUC of approximately 0.508 demonstrates weak discrimination across the three classes [54]. Fourth, the analysis identifies statistical and

predictive relationships but does not establish causal relationships between individual risk factors and disruptions. Finally, predictive performance may change over time as logistics conditions, disruption patterns, and industry-specific characteristics evolve [55].

Future research should address the class imbalance using techniques such as Synthetic Minority Over-Sampling Technique (SMOTE), adaptive resampling, class-weight optimization, and cost-sensitive learning [56], [57]. More advanced models, including XGBoost, LightGBM, neural networks, recurrent neural networks, and ensemble stacking, should be tested to determine whether they improve multiclass discrimination and ROC-AUC performance. Explainable-AI techniques such as SHAP and Local Interpretable Model-Agnostic Explanations (LIME) could improve managerial transparency and confidence in individual risk predictions [58]. Future work should also incorporate streaming data from GPS, IoT sensors, weather, traffic, port operations, and supplier monitoring to support dynamic rather than primarily static risk prediction [59].

The framework should also be tested using data from other U.S. regions and critical sectors such as transportation, manufacturing, healthcare, food distribution, and energy to improve generalizability [60], [61]. A further step would be to develop a real-time decision-support system in which predicted risk levels trigger alerts and recommended mitigation measures such as route changes, inventory adjustments, supplier escalation, or resource reallocation. Such development should be accompanied by validation of data governance, privacy, cybersecurity, communication reliability, model drift, and human oversight before operational deployment.

## 5. Conclusion

This study explored the feasibility of using AI-powered predictive analytics to support proactive supply-chain risk management and resilience in critical U.S. industries using a logistics and supply-chain dataset representing a Southern California network. The analysis shows that supply-chain risk is multidimensional and is associated with transportation, inventory, supplier, environmental, operational, demand, and human-related conditions. The dataset provides several useful predictive outcomes, including Disruption Likelihood Score, Delay Probability, Risk Classification, and Delivery Time Deviation. A substantial imbalance was observed among the risk classes, with most observations classified as High Risk. Disruption Likelihood Score clearly distinguished the three risk categories, whereas Delay Probability overlapped considerably across them, indicating that different risk indicators provide different levels of predictive information. Feature-importance analysis identified lead time, warehouse inventory, ETA variation, order fulfillment, cargo condition, driver behavior, historical demand, fatigue monitoring, and equipment availability as influential factors. Random Forest, HistGradient Boosting, and Support Vector Machine achieved stronger aggregate performance than Logistic Regression and Decision Tree, but the Random Forest confusion matrix and macro-AUC of approximately 0.508 revealed major weaknesses in minority-class recognition and multiclass discrimination. The findings therefore show that overall accuracy alone is insufficient for evaluating AI-based supply-chain risk prediction under severe class imbalance. AI-driven predictive analytics can contribute to early-warning systems and proactive decision-making, but reliable implementation requires careful model evaluation, class-balance management, data-quality assurance, interpretability, and expert oversight. The proposed framework provides a structured foundation linking multidimensional risk data, predictive analytics, early-warning insights, proactive mitigation, and supply-chain resilience.

## REFERENCES

[1] M. Ali, M. S. Hossain, M. W. Rahman, and M. S. Hossain, “Leveraging business analytics to enhance supply chain resilience and reduce disruptions in critical US industries,” Journal of Business and Management Studies, vol. 4, no. 4, pp. 239–263, 2022.

[2] M. R. Hasan, M. A. Rahman, C. A. H. Gomes, F. N. Nitu, C. A. Gomes, M. R. Islam, and R. E. R. Shawon, “Building robust AI and machine learning models for supplier risk management: A data-driven strategy for enhancing supply chain resilience in the USA,” Advances in Consumer Research, vol. 2, no. 4, pp.

1152–1171, 2025.
[3] B. Sarker, K. P. Mishu, M. T. Ahmed, N. K. Papri, A. Sarker, and M. Y. Ahmad, "An AI-driven optimization and risk mitigation framework to strengthen US supply chain resilience and national economic security," American Journal of Engineering, Mechanics and Architecture, vol. 1, no. 10, pp. 418–427, 2023.
[4] A. Salehi, A. Babaei, and H. Hamidi, "AI-driven strategies for supply chain resilience: A review of challenges and solutions during pandemics," Int. J. Eng., vol. 39, no. 3, pp. 585–605, 2026.
[5] M. Bahrami and S. Shokouhyar, "The role of big data analytics capabilities in bolstering supply chain resilience and firm performance: A dynamic capability view," Information Technology & People, vol. 35, no. 5, pp. 1621–1651, 2022.
[6] M. S. Golan, B. D. Trump, J. C. Cegan, and I. Linkov, "Supply chain resilience for vaccines: Review of modeling approaches in the context of the COVID-19 pandemic," Industrial Management & Data Systems, vol. 121, no. 7, pp. 1723–1748, 2021.
[7] M. Munir, M. S. S. Jajja, and K. A. Chatha, "Capabilities for enhancing supply chain resilience and responsiveness in the COVID-19 pandemic: Exploring the role of improvisation, anticipation, and data analytics capabilities," International Journal of Operations & Production Management, vol. 42, no. 10, pp. 1576–1604, 2022.
[8] A. Rashid and R. Rasheed, "Supply chain cognitive resilience through AI-powered risk mitigation and autonomous resilience," Supply Chain Management: An International Journal, vol. 31, no. 2, pp. 222–237, 2026.
[9] F. Nsama, "Strategic development of AI-driven supply chain resilience frameworks for critical US sectors," Communication in Physical Sciences, vol. 12, no. 5, pp. 1538–1560, 2025.
[10] V. Diyora, "AI-driven data analytics and risk intelligence for US critical infrastructure and business resilience," 2024.
[11] A. Sheikh and S. Sajja, "AI-driven risk mitigation in global supply chains to build resilience for US competitiveness: A conceptual review and framework-based study," Journal of Global Economics, Management and Business Research, vol. 18, no. 1, pp. 198–206, 2026.
[12] H. Ahmed, "AI-driven predictive supply chain resilience framework for US manufacturing systems: An empirical analysis," American Journal of Data Science and Analytics, vol. 7, no. 4, pp. 45–80, 2026.
[13] A. Rashid and R. Rasheed, "AI-powered risk mitigation and organizational sensemaking for supply chain resilience," Supply Chain Management: An International Journal, vol. 31, no. 4, pp. 610–625, 2026.
[14] D. Ivanov and A. Dolgui, "A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0," Production Planning & Control, vol. 32, no. 9, pp. 775–788, 2021.
[15] H. Liu and S. Wei, "Leveraging supply chain disruption orientation for resilience: The roles of supply chain risk management practices and analytics capability," International Journal of Physical Distribution & Logistics Management, vol. 52, nos. 9–10, pp. 771–790, 2022.
[16] R. S. Khan, M. R. M. Sirazy, R. Das, and S. Rahman, "An AI and ML-enabled framework for proactive risk mitigation and resilience optimization in global supply chains during national emergencies," Sage Science Review of Applied Machine Learning, vol. 5, no. 2, pp. 127–144, 2022.
[17] S. R. Katragadda, "AI-driven resilient supply chain architectures: Machine learning frameworks for risk anticipation, disruption mitigation, and adaptive decision-making," Journal of Business and Management Studies, vol. 8, no. 3, pp. 38–50, 2026.
[18] S. O. Usoro, I. A. Omogiate, R. Z. Felix, and D. A. Galadima, "AI-driven risk forecasting for strengthening the United States food supply chain resilience: Case study: A national AI-enabled food supply chain risk forecasting framework and a data analytics approach to predicting disruptions," International Journal of Science and Research Archive, vol. 18, no. 1, p. 744, 2026.
[19] K. P. Mishu, A. Sarker, N. K. Papri, M. T. Ahmed, and B. Sarker, "An AI- and machine-learning data-driven predictive analytics framework for enhancing resilience and sustainable US supply chains systems for manufacturing and resource efficiency," Journal of Computer Science and Information Technology, vol. 3, no. 1, pp. 1–11, 2025.
[20] K. P. Mishu, A. Sarker, N. K. Papri, M. T. Ahmed, and B. Sarker, "An AI- and machine-learning data-

driven predictive analytics framework for enhancing resilience and sustainable US supply chains systems for manufacturing and resource efficiency," Journal of Computer Science and Information Technology, vol. 3, no. 1, pp. 1–11, 2025.
[21] C. O. Nwamekwe and N. C. Igbokwe, "Supply chain risk management: Leveraging AI for risk identification, mitigation, and resilience planning," International Journal of Industrial Engineering, Technology & Operations Management, 2024.
[22] I. Zerine, M. S. Islam, M. Y. Ahmad, M. M. Islam, and Y. A. Biswas, "AI-driven supply chain resilience: Integrating reinforcement learning and predictive analytics for proactive disruption management," Business and Social Sciences, vol. 1, no. 1, pp. 1–12, 2023.
[23] R. Ejjami and K. Boussalham, "Resilient supply chains in Industry 5.0: Leveraging AI for predictive maintenance and risk mitigation," International Journal for Multidisciplinary Research, vol. 6, no. 4, p. 25116, 2024.
[24] A. A. Ogunnubi, C. F. Davis, and F. O. Agbadzete, "AI-driven financial analytics for supply chain resilience in US manufacturing," 2025.
[25] T. Pujiati, M. Kamil, N. Silawati, and R. S. Ikhsan, "Integrating AI-driven predictive analytics and smart contracts for data-driven supply chain risk management," ADI Journal on Recent Innovation (AJRI), vol. 7, no. 1, pp. 50–61, 2025.
[26] H. Wu, J. Liu, and B. Liang, "AI-driven supply chain transformation in Industry 5.0: Enhancing resilience and sustainability," Journal of the Knowledge Economy, vol. 16, no. 1, pp. 3826–3868, 2025.
[27] A. Singh and G. Singh, "Leveraging AI-powered predictive analytics for proactive risk management and business resilience in SMEs: Harnessing predictive insights to strengthen decision-making in SMEs," in Incentives and Benefits for Adopting Green Entrepreneurship Practices, pp. 265–294, 2026.
[28] A. O. Adewusi, A. M. Komolafe, E. Ejairu, I. A. Aderotoye, O. O. Abiona, and O. C. Oyeniran, "The role of predictive analytics in optimizing supply chain resilience: A review of techniques and case studies," International Journal of Management & Entrepreneurship Research, vol. 6, no. 3, pp. 815–837, 2024.
[29] U. Iqbal, "AI-powered supplier risk intelligence: Predicting financial and geopolitical supply chain disruptions in US critical industries," Journal of Engineering and Computational Intelligence Review, vol. 3, no. 2, pp. 173–193, 2025.
[30] R. Ejjami and K. Boussalham, "Industry 5.0 in manufacturing: Enhancing resilience and responsibility through AI-driven predictive maintenance, quality control, and supply chain optimization," International Journal for Multidisciplinary Research, vol. 6, no. 4, pp. 1–31, 2024.
[31] K. Beta, S. S. Nagaraj, and T. D. Weerasinghe, "The role of artificial intelligence on supply chain resilience," Journal of Enterprise Information Management, vol. 38, no. 3, pp. 950–973, 2025.
[32] R. Deora, "Leveraging AI for supply chain resilience: A deep learning approach to risk management," Nanotechnology Perceptions, 2024.
[33] A. Aljohani, "Predictive analytics and machine learning for real-time supply chain risk mitigation and agility," Sustainability, vol. 15, no. 20, p. 15088, 2023.
[34] N. Chukwu, S. Yufenyuy, E. Ejiofor, D. Ekweli, O. Ogunleye, T. Clement, et al., "Resilient chain: AI-enhanced supply chain security and efficiency integration," International Journal of Scientific and Management Research, vol. 7, no. 3, pp. 46–65, 2024.
[35] W. O. Nduka, S. S. Abba, A. O. Ariyo, A. O. Olisa, and A. M. Ogunmolu, "AI-driven predictive resilience: Integrating impact forecasting, governance, and proactive mitigation in networks," Journal of Engineering Research and Reports, vol. 28, no. 3, pp. 75–90, 2026.
[36] M. I. Faruk, F. W. Plabon, U. S. Saha, and M. D. Hossain, "AI-driven project risk management: Leveraging artificial intelligence to predict, mitigate, and manage project risks in critical infrastructure and national security projects," Journal of Computer Science and Technology Studies, vol. 7, no. 6, pp. 123–137, 2025.
[37] U. Iqbal, "AI-driven predictive maintenance for US smart manufacturing: Deep learning models for equipment failure prediction and operational resilience," Journal of Engineering and Computational Intelligence Review, vol. 3, no. 1, pp. 114–138, 2025.
[38] J. Rane, R. A. Chaudhari, and N. Rane, "Artificial intelligence and machine learning for supply chain

resilience: Risk assessment and decision making in manufacturing Industry 4.0 and 5.0," Artificial Intelligence and Machine Learning for Supply Chain Resilience: Risk Assessment and Decision Making in Manufacturing Industry, vol. 4, 2025.

[39] M. S. Uddin, M. S. Sikder, M. M. Anwar, and F. Hossain, "AI-driven cybersecurity and big data-enabled MIS frameworks: Strengthening supply chain integrity, energy resilience, and critical infrastructure protection," Journal of Computer Science and Technology Studies, vol. 7, no. 9, pp. 223–232, 2025.

[40] M. R. I. Khan, M. S. Tahsin, A. Al Mamun, K. Al-Samad, H. Ahmed, S. A. Al Wahid, et al., "Multimodal machine learning for proactive supply chain risk management: An AI-driven framework integrating sensor, operational, and external intelligence," Journal of Computer Science and Technology Studies, vol. 7, no. 9, pp. 657–673, 2025.

[41] I. Jahan, M. H. O. Rashid, S. Nashid, and K. N. Pinky, "AI-driven predictive risk modeling to improve cost efficiency and system resilience in US healthcare infrastructure," Power System Protection and Control, vol. 54, no. 2, pp. 48–61, 2026.

[42] M. Riad, M. Naimi, and C. Okar, "Enhancing supply chain resilience through artificial intelligence: Developing a comprehensive conceptual framework for AI implementation and supply chain optimization," Logistics, vol. 8, no. 4, p. 111, 2024.

[43] S. M. Islam, M. S. Bari, A. Sarkar, A. Obaidur, R. Khan, and R. Paul, "AI-driven threat intelligence: Transforming cybersecurity for proactive risk management in critical sectors," International Journal of Computer Science and Information Technology, vol. 16, no. 5, pp. 125–131, 2024.

[44] A. Belhadi, V. Mani, S. S. Kamble, S. A. R. Khan, and S. Verma, "Artificial intelligence-driven innovation for enhancing supply chain resilience and performance under the effect of supply chain dynamism: An empirical investigation," Annals of Operations Research, vol. 333, no. 2, pp. 627–652, 2024.

[45] R. U. Attah, B. M. P. Garba, I. Gil-Ozoudeh, and O. Iwuanyanwu, "Enhancing supply chain resilience through artificial intelligence: Analyzing problem-solving approaches in logistics management," International Journal of Management & Entrepreneurship Research, vol. 5, no. 12, pp. 3248–3265, 2024.

[46] L. W. Wong, G. W. H. Tan, K. B. Ooi, B. Lin, and Y. K. Dwivedi, "Artificial intelligence-driven risk management for enhancing supply chain agility: A deep-learning-based dual-stage PLS-SEM-ANN analysis," International Journal of Production Research, vol. 62, no. 15, pp. 5535–5555, 2024.

[47] O. Ayeni, "Resilient project leadership: Mitigating uncertainty, regulatory constraints, and global supply chain disruptions," International Research Journal of Modernization in Engineering Technology and Science, vol. 7, no. 3, pp. 216–235, 2025.

[48] S. K. Anumula, V. Krishnapillai, and D. K. Rai, "Optimizing supply chain management with AI-powered predictive analytics," Journal of International Commercial Law and Technology, vol. 6, pp. 244–252, 2025.

[49] N. Qahtani, "Enhancing US critical mineral supply chain resilience through predictive analytics and risk control engineering," 2026.

[50] B. Sarker, "Leveraging machine learning for supply chain disruption prediction and risk assessment in US industries," International Journal of Applied Mathematics, vol. 38, no. 10s, pp. 2382–2410, 2025.

[51] C. Smyth, D. Dennehy, S. Fosso Wamba, M. Scott, and A. Harfouche, "Artificial intelligence and prescriptive analytics for supply chain resilience: A systematic literature review and research agenda," International Journal of Production Research, vol. 62, no. 23, pp. 8537–8561, 2024.

[52] E. Hiljemark and D. Nika, "AI innovations in risk management—A case study of Volvo AB's supply chain resilience," 2024.

[53] S. Modgil, R. K. Singh, and C. Hannibal, "Artificial intelligence for supply chain resilience: Learning from COVID-19," The International Journal of Logistics Management, vol. 33, no. 4, pp. 1246–1268, 2022.

[54] F. Ahmed, M. M. Rahman, and S. Akter, "Reducing supply chain disruptions, costs, and waste using AI and blockchain to strengthen US economic resilience," Journal of Tianjin University Science and Technology, vol. 58, pp. 470–494, 2025.

[55] H. Kagalwala, G. V. Radhakrishnan, I. A. Mohammed, R. R. Kothinti, and N. Kulkarni, "Predictive analytics in supply chain management: The role of AI and machine learning in demand forecasting," Advances in Consumer Research, vol. 2, pp. 142–149, 2025.

[56] J. Rane, R. A. Chaudhari, and N. L. Rane, Enhancing Sustainable Supply Chain Resilience Through Artificial Intelligence and Machine Learning: Industry 4.0 and Industry 5.0 in Manufacturing. Deep Science Publishing, 2025.

[57] M. I. Khan, R. Lama, K. S. Islam, M. R. Karim, and B. Tamang, "AI-driven risk management & optimization in healthcare supply chain: A machine learning approach," Journal of Artificial Intelligence General Science (JAIGS), vol. 8, no. 2, pp. 149–160, 2025.

[58] A. Rashid, R. Rasheed, A. H. Ngah, and N. A. Amirah, "Unleashing the power of cloud adoption and artificial intelligence in optimizing resilience and sustainable manufacturing supply chain in the USA," Journal of Manufacturing Technology Management, vol. 35, no. 7, pp. 1329–1353, 2024.

[59] Z. Zong and Y. Guan, "AI-driven intelligent data analytics and predictive analysis in Industry 4.0: Transforming knowledge, innovation, and efficiency," Journal of the Knowledge Economy, vol. 16, no. 1, pp. 864–903, 2025.

[60] S. Modgil, S. Gupta, R. Stekelorum, and I. Laguir, "AI technologies and their impact on supply chain resilience during COVID-19," International Journal of Physical Distribution & Logistics Management, vol. 52, no. 2, pp. 130–149, 2022.

[61] S. Gupta, S. Bag, S. Modgil, A. B. L. de Sousa Jabbour, and A. Kumar, "Examining the influence of big data analytics and additive manufacturing on supply chain risk control and resilience: An empirical study," Computers & Industrial Engineering, vol. 172, p. 108629, 2022.

[62] "Logistics and Supply Chain Dataset," Kaggle. [Online]. Available: https://www.kaggle.com/datasets/datasetengineer/logistics-and-supply-chain-dataset.