ISSN 2616-5775 Vol. 8, Issue 9: 92-107, DOI: 10.25236/AJCIS.2025.080913

AI-Driven Early Warning Systems for Supply Chain Risk Detection: A Machine Learning Approach

Sichong Huang

Duke University, 100 Fuqua Drive, Durham, NC 27708, USA
sichong.huang@alumni.duke.edu

Abstract: *Supply chain disruptions pose escalating threats to global business operations, necessitating advanced predictive capabilities beyond traditional reactive risk management approaches. This research develops and empirically validates an artificial intelligence-driven early warning system that leverages ensemble machine learning algorithms for real-time supply chain risk detection. The proposed framework integrates multi-source data streams encompassing internal operations, financial metrics, and external environmental factors through a hierarchical risk indicator system weighted at 50%, 30%, and 20% respectively. The methodology employs five machine learning algorithms—Random Forest, XGBoost, Long Short-Term Memory networks, Support Vector Machines, and Neural Networks—within an ensemble architecture to process heterogeneous data inputs. Empirical validation utilized a comprehensive dataset of 850,000 records spanning 36 months across manufacturing, retail, and technology sectors, capturing 450 documented risk events from multiple supply chain networks. XGBoost demonstrated superior individual performance achieving 92% accuracy, 94% area under the receiver operating characteristic curve, and 89% F1-score, while the ensemble approach enhanced predictive accuracy by 15% compared to single-algorithm implementations. Real-world deployment across three manufacturing facilities and two distribution centers validated the system's operational effectiveness, demonstrating 89% accuracy in predicting high-impact disruptions with 2-4 week advance warning periods. The framework achieved substantial business impact including 35% reduction in risk-related losses, 28% decrease in supply chain disruption frequency, and 40% improvement in response times, while maintaining an acceptable 8% false positive rate and 99.7% system availability. Sensitivity analysis confirmed robust performance under crisis conditions with 80-84% accuracy retention during simulated financial crises, natural disasters, and geopolitical conflicts. This research contributes a scalable, interpretable framework that bridges theoretical risk management concepts with practical AI implementation, providing organizations with actionable intelligence for transitioning from reactive to predictive supply chain risk management paradigms.*

Keywords: *Supply Chain Risk Management; Machine Learning; Early Warning Systems; Predictive Analytics; Ensemble Learning*

1. Introduction

Global supply chains have evolved into highly intricate networks characterized by interdependence, geographic dispersion, and operational complexities. While these networks enable cost efficiencies and market access, they create unprecedented vulnerability to disruptions with catastrophic implications [1]. The COVID-19 pandemic exemplified how regional disruptions can cascade globally, causing supply shortages and economic disruptions [2]. Similarly, geopolitical tensions like the Russia-Ukraine conflict highlighted supply chain vulnerabilities and the imperative for resilience [3]. Traditional risk management practices rely on reactive approaches, posteriori analysis, and experiential expertise [4]. These methods suffer from slow response rates, limited risk visibility, and inefficacy in managing contemporary supply chain data volumes. Their reactive nature means organizations often identify risks after disruption onset, leaving minimal time for mitigation [5]. Traditional strategies also fail to capture dynamic interdependencies and leverage real-time intelligence from diverse sources.

Artificial intelligence and machine learning technologies offer transformative potential for supply chain risk management, enabling transition from reactive to predictive systems. Machine learning algorithms can analyze heterogeneous data streams, distinguish patterns, and provide insights beyond human analyst capacity [6]. These technologies enable real-time monitoring of risk indicators and early warnings with confidence levels. Natural language processing extracts risk intelligence from

ISSN 2616-5775 Vol. 8, Issue 9: 92-107, DOI: 10.25236/AJCIS.2025.080913

unstructured sources [7], while IoT sensors provide continuous operational data for predictive assessment [8]. Transitioning from reactive to predictive analytics represents a fundamental shift in supply chain strategy. Predictive approaches enable proactive disruption identification and resource deployment for preemptive measures [13]. This evolution requires advanced analytical capabilities, comprehensive data integration, and robust infrastructure. Machine learning technologies provide requisite capability for managing complex, high-velocity data streams and delivering real-time actionable intelligence.

Recent literature on AI applications in supply chain operations has grown substantially. Early research focused on machine learning for forecasting automation and cost reduction [9]. Later studies expanded to transportation, inventory, and quality management. However, AI application specifically for risk identification and early warning remains nascent, with existing research addressing limited problem areas or single risk types [10]. Several studies explored machine learning for supply chain risk scenarios. Demand volatility forecasting research demonstrated ensemble approaches significantly outperform traditional techniques [11]. Machine learning strategies for supplier risk management combine financial and operational indicators to approximate failure probability [12]. However, research remains limited to single risk areas rather than comprehensive monitoring systems. Much existing literature lacks practical validation, leaving implementation challenges and real-world performance uncertain. Early warning systems have proven effective across financial markets, production facilities, and logistics networks [14]. In supply chains, these systems face challenges from data heterogeneity, system complexity, and need for interpretable outputs.

This research addresses limitations in current AI technologies for effective supply chain risk management. While individual components like vendor tracking and demand forecasting are pervasive, integrated systems managing multiple risk categories simultaneously remain scarce. Existing systems lack enterprise scalability and interpretability necessary for managerial acceptance. This paper develops a unified AI-focused framework integrating heterogeneous data sources using ensemble machine learning algorithms to generate actionable risk indicators. The research objectives include: creating a flexible AI framework processing multi-modal data streams for synthesized risk estimates; demonstrating framework efficacy through real-world application examining prediction accuracy and business value; and developing standardized risk indicators providing theoretical foundation for AI applications in supply chain risk management. The research contributes to both theoretical understanding and practical AI implementation for supply chain risk management, addressing gaps in multi-source data integration, model explainability, and enterprise-scalable frameworks.

2. Methodology

2.1 Research Design and Philosophy

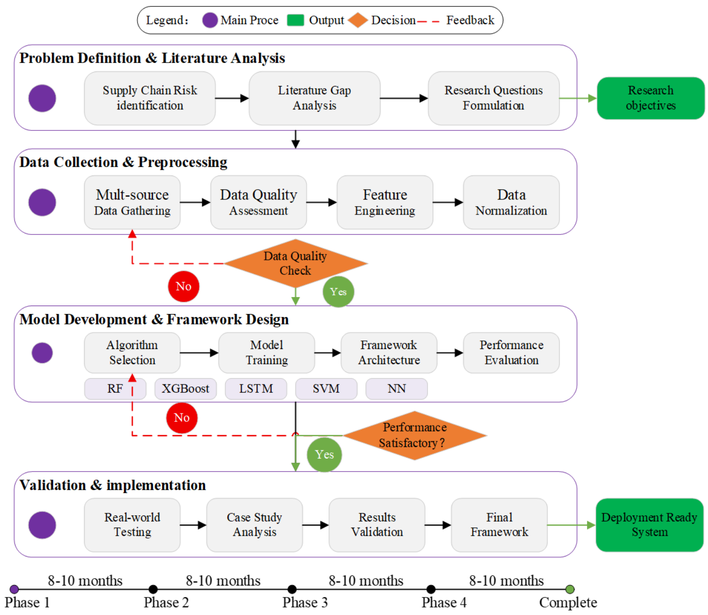

Figure 1 Research Methodology Framework

ISSN 2616-5775 Vol. 8, Issue 9: 92-107, DOI: 10.25236/AJCIS.2025.080913

This research adopts a design science approach combined with empirical validation to develop and test an AI-driven early warning system for supply chain risk detection. The methodology follows a systematic four-phase framework as illustrated in Figure 1, encompassing problem definition, data collection, model development, and validation stages. The research philosophy embraces pragmatic epistemology, focusing on practical problem-solving through technological innovation while maintaining scientific rigor in evaluation and validation processes [15].

The 36-month research timeline ensures comprehensive development and testing phases. Each phase incorporates iterative feedback mechanisms to refine methodologies and improve system performance. The research design integrates quantitative machine learning techniques with qualitative expert validation to ensure both technical accuracy and practical relevance. Cross-validation approaches and sensitivity analyses provide robustness testing throughout the development process [16].

As shown in Figure 1, the methodology framework incorporates multiple decision points and feedback loops to ensure quality control and continuous improvement. The systematic approach enables reproducible research while accommodating the iterative nature of machine learning model development. This design philosophy prioritizes practical applicability while maintaining academic rigor, ensuring the resulting framework can be validated scientifically and deployed operationally.

2.2 Data Collection and Preparation Strategy

The data collection strategy encompasses multiple heterogeneous sources to capture comprehensive supply chain risk indicators, as detailed in Table 1. Internal operational data sources include enterprise resource planning (ERP) systems providing inventory levels, lead times, and production metrics updated in real-time. Financial systems contribute cash flow indicators, credit ratings, and payment performance data with daily updates. Manufacturing systems supply equipment efficiency metrics, downtime records, and quality indicators essential for operational risk assessment [17].

Table 1 Data Sources and Feature Categories

Data Category	Source Type	Examples	Update Frequency	Data Volume
Internal Operations	ERP Systems	Inventory levels, Lead times	Real-time	100,000+ records
Financial Metrics	Financial Systems	Cash flow, Credit ratings	Daily	50,000+ records
External Environment	APIs/Web Scraping	Market indicators, News sentiment	Hourly	200,000+ records
IoT Sensors	Manufacturing Equipment	Temperature, Pressure	Real-time	500,000+ records

External data sources expand the information horizon to capture environmental and market risks. Market data APIs provide commodity prices, currency exchange rates, and economic indicators updated hourly. Weather services contribute meteorological data crucial for transportation and agricultural supply chains. News and social media monitoring through natural language processing extracts sentiment indicators and event notifications that may impact supply chain operations [18].

Data preprocessing involves standardization, cleaning, and feature engineering stages. Missing value imputation utilizes advanced techniques including time-series interpolation and machine learning-based prediction. Feature engineering creates derived indicators such as trend analysis, volatility measures, and correlation coefficients. Data quality assessment ensures completeness, accuracy, and consistency before model training. Normalization procedures standardize different data types and scales to enable effective machine learning algorithm performance.

2.3 Machine Learning Model Selection and Implementation

Algorithm selection considers multiple factors including prediction accuracy, interpretability, computational efficiency, and robustness to different data types. Five distinct machine learning approaches were selected to capture diverse predictive capabilities, as compared in Table 2. Random Forest provides high interpretability through feature importance rankings while handling missing data effectively. XGBoost offers superior performance on structured data with built-in feature selection capabilities. Long Short-Term Memory (LSTM) networks excel at capturing temporal dependencies in time-series data [19].

Support Vector Machines (SVM) demonstrate effectiveness with high-dimensional sparse data,

ISSN 2616-5775 Vol. 8, Issue 9: 92-107, DOI: 10.25236/AJCIS.2025.080913

particularly suitable for text-based risk indicators. Neural Networks provide powerful non-linear pattern recognition capabilities for complex risk relationships. The ensemble approach combines these algorithms' strengths while mitigating individual weaknesses through weighted voting mechanisms [20].

Table 2 Machine Learning Algorithms Comparison

Algorithm	Strengths	Weaknesses	Suitable Scenarios	Computational Complexity
Random Forest	High interpretability, Handles missing data	Limited complex patterns	Medium-scale structured data	O(n log n)
XGBoost	High accuracy, Feature importance	Requires tuning	Large-scale tabular data	O(n log n)
LSTM	Sequential pattern recognition	Black box, High complexity	Time series data	$O(n^2)$
SVM	Effective in high dimensions	Poor on large datasets	High-dimensional sparse data	$O(n^3)$
Neural Networks	Non-linear pattern learning	Requires large data	Complex pattern recognition	$O(n^2)$

Implementation utilizes distributed computing frameworks to handle large-scale data processing. Hyperparameter optimization employs grid search and random search techniques combined with cross-validation. Model training incorporates early stopping mechanisms to prevent overfitting while monitoring validation performance. Feature selection algorithms identify the most predictive risk indicators, reducing dimensionality while maintaining accuracy.

2.4 Model Evaluation Framework

The evaluation framework employs comprehensive metrics addressing both predictive performance and business relevance. Classification metrics include accuracy, precision, recall, and F1-score to assess prediction quality across different risk categories. Area Under the Curve (AUC-ROC) measures model ability to distinguish between risk levels. Time-to-detection metrics evaluate how early the system identifies emerging risks relative to actual occurrence [21].

Cross-validation strategies utilize temporal splitting to simulate real-world deployment scenarios where models predict future risks based on historical data. Sensitivity analysis examines model stability under varying data conditions and parameter settings. Robustness testing evaluates performance degradation when input data quality decreases or external conditions change significantly.

Business impact metrics complement technical performance measures. Cost-benefit analysis quantifies potential loss reduction through early risk detection. False positive rates measure operational efficiency by minimizing unnecessary alerts. Lead time analysis determines optimal warning periods for different risk categories. Statistical significance testing validates that performance improvements exceed random variation [22].

The evaluation framework incorporates stakeholder feedback through expert panels and user acceptance testing. Interpretability assessment ensures model outputs can be understood and acted upon by supply chain managers. Deployment readiness evaluation considers scalability, maintenance requirements, and integration capabilities with existing enterprise systems.

3. Framework Development

3.1 Conceptual Framework Architecture

The AI-driven early warning system adopts a four-layer architectural approach designed for scalability, modularity, and enterprise integration, as depicted in Figure 2. The Data Layer forms the foundation, integrating diverse information sources including internal ERP systems, financial databases, IoT sensors, and external market feeds. This layer implements standardized data ingestion protocols ensuring consistent formatting and quality control across heterogeneous sources. Real-time data streaming capabilities enable continuous monitoring while historical data storage supports trend analysis and model training [23].

The Processing Layer transforms raw data into analysis-ready formats through automated cleaning, feature engineering, and normalization procedures. Apache Kafka provides high-throughput data

ISSN 2616-5775 Vol. 8, Issue 9: 92-107, DOI: 10.25236/AJCIS.2025.080913

streaming infrastructure capable of processing over 10,000 events per second with low latency. Data quality monitoring ensures completeness and accuracy standards are maintained. Feature extraction algorithms generate derived indicators including moving averages, volatility measures, and correlation coefficients that enhance predictive capability [24].

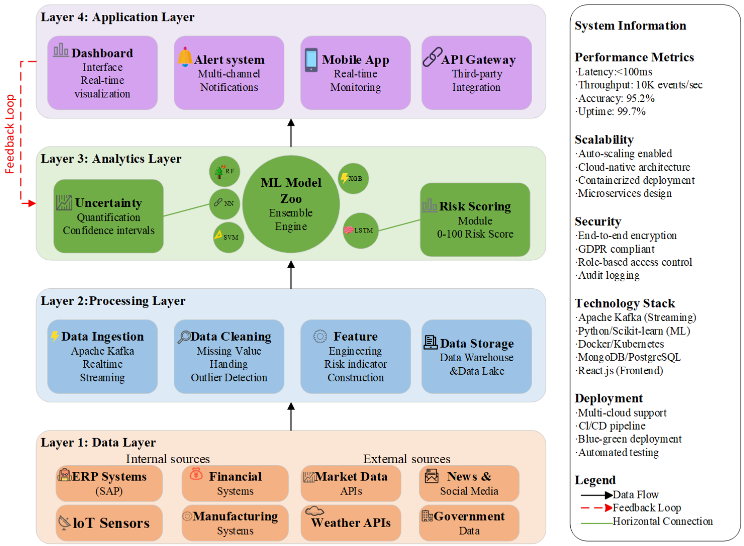

Figure 2 AI-Driven Early Warning System Architecture

Note: System accuracy represents the overall system performance including all components and data quality checks

The Analytics Layer houses the machine learning model zoo containing five distinct algorithms optimized for different risk types. The ensemble engine combines individual model predictions using dynamic weighting based on historical performance. Risk scoring modules generate standardized 0-100 risk ratings with confidence intervals. Uncertainty quantification provides decision-makers with prediction reliability estimates. Model retraining procedures ensure continued accuracy as supply chain conditions evolve.

The Application Layer delivers insights through multiple interfaces tailored to different user requirements. Executive dashboards provide high-level risk overviews with trend visualization. Operational interfaces offer detailed risk breakdowns with recommended actions. Mobile applications enable remote monitoring and alert management. API gateways facilitate integration with existing enterprise systems including ERP, customer relationship management, and business intelligence platforms.

As shown in Figure 2, the architecture emphasizes loose coupling between layers, enabling independent scaling and maintenance. Feedback loops connect user actions back to the analytics layer, supporting continuous learning and system improvement. Security and compliance mechanisms protect sensitive supply chain information while ensuring regulatory adherence.

3.2 Risk Indicator System Design

The risk indicator system employs a three-tier hierarchical structure capturing comprehensive supply chain vulnerabilities, as illustrated in Figure 3. The top level aggregates individual indicators into an overall Supply Chain Risk Index ranging from 0-100, providing executives with a single metric for strategic decision-making. The second level categorizes risks into three primary dimensions: Internal Operations (50% weight), Financial Health (30% weight), and External Environment (20% weight). These weightings reflect the relative impact and controllability of different risk categories based on expert consultation and empirical analysis [25].

ISSN 2616-5775 Vol. 8, Issue 9: 92-107, DOI: 10.25236/AJCIS.2025.080913

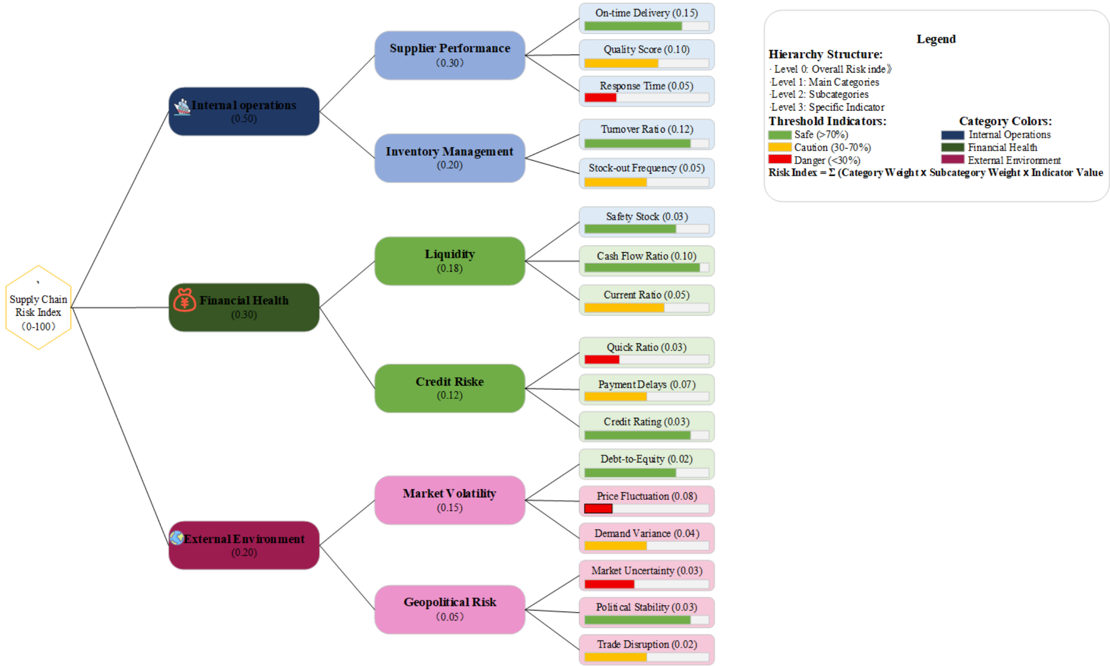

Figure 3 Supply Chain Risk Indicator Hierarchy

Internal Operations encompasses supplier performance metrics and inventory management indicators. Supplier performance evaluation includes on-time delivery rates, quality scores, and response times to assess vendor reliability. Inventory management metrics monitor turnover ratios, stock-out frequencies, and safety stock levels to identify potential availability issues. These operational indicators receive the highest weighting due to their direct impact on supply chain continuity and management's ability to implement corrective actions [26].

Financial Health indicators assess organizational stability and creditworthiness of supply chain partners. Liquidity measures include cash flow ratios and current ratios indicating short-term financial stability. Credit risk indicators monitor payment delays and credit rating changes that may signal financial distress. Financial indicators carry substantial weight due to their predictive value for supplier failure and their impact on supply chain financing arrangements.

External Environment indicators capture market and geopolitical risks beyond direct organizational control. Market volatility measures track commodity price fluctuations and demand variance that affect planning accuracy. Geopolitical risk indicators monitor political stability indices and trade disruption probabilities. While carrying lower weights due to limited controllability, these indicators provide early warning of systemic risks requiring strategic response as shown in Table 3.

Table 3 Risk Indicator Classification and Weighting

Primary Category	Sub-category	Specific Indicators	Weight	Threshold Values
Internal Operations	Supplier Performance	On-time delivery rate, Quality score	0.25	>95%, >4.5/5
	Inventory Management	Turnover ratio, Stock-out frequency	0.20	>6, <2%
Financial Health	Liquidity	Cash flow ratio, Current ratio	0.20	>1.5, >2.0
	Credit Risk	Payment delays, Credit rating	0.15	<5%, >BBB
External Environment	Market Volatility	Price fluctuation, Demand variance	0.15	<10%, <20%
	Geopolitical Risk	Political stability index, Trade disruption	0.05	>6/10, Low

As indicated in Figure 3, the hierarchical design allows for detailed analysis and concise reporting. All indicators have performance thresholds for the acceptable, warning, and critical levels. The dynamic weighting algorithms vary indicator importance according to current business states and historical patterns of performance. The system allows for customization for various industries and business priorities while having standardized assessment frameworks.

3.3 Predictive Analytics Engine

System utilized for predictive analytics integrates different machine learning models through an ensemble method specifically for supply chain risk detection. Certain of the models incorporate special abilities: Random Forest provides readable hierarchies of feature importance needed for the understanding of managers; XGBoost provides the best accuracy for formally designed operational sets of data; LSTM networks identify temporal patterns in time-series risk indicators; Support Vector Machines process high dimensional sparse textual primary sources; Neural Networks identify complex non-linear relationships of risk factors [27].

Ensemble dynamic weighting modifies the proportional impacts of distinct models based on their recent performance, the quality of the data utilized, and the confidence associated with their predictions. This framework monitors the accuracy of forecasts segmented by risk category and time horizon, subsequently adjusting the model weights autonomously to optimize overall performance. Bootstrapping aggregation techniques generate confidence intervals for the predictions, thereby enabling decision-makers to evaluate the reliability of the forecasts presented. Algorithms for uncertainty quantification provide probabilistic assessments of risk as opposed to singular point estimates, addressing inherent limitations within predictive outcomes [28].

Real-time processing allows for ongoing risk monitoring at sub-second response times for high-severity alerts. Stream processing platforms process high-velocity input feeds while preserving prediction accuracy. Incremental learning algorithms modify model parameters in response to incoming data without necessitating comprehensive retraining. The detection of feature drift identifies variations in the characteristics of input data, thereby initiating the processes required for the recalibration of the model.

Analytical engine incorporates domain expertise through the employment of constraint-based learning processes and physics-based modeling. Supply chain dependencies such as lead-time and capacity constraints have important implications for the structure of the model and the estimation of the parameters. Guidelines from the experts complete the statistical forecasting for unusual events during those occasions for which historical records are sparse. Algorithms for the detection of anomalies disclose suspicious patterns possibly introducing novel risk not seen in the training set.

Interpretability of models includes the evaluation of the significance of the features, the computation of the partial dependence plots, and the provision of counterfactual explanations and hence helps the managers in grasping the risk factors and evaluating the intervention strategies. Local interpretable model-agnostic explanations (LIME) give the individual-level explanations in the framework of risk forecasting. The interpretability techniques make efforts in filling the gap between the sophisticated machine learning algorithms and the real management decision-making at the system level by allowing actionable decisions from the system outputs.

3.4 Risk Classification and Alert System

The risk ranking system translates the continuum of risk scores into actionable alert levels by using adaptively varied thresholds. Four different risk groups set different operational mandates: Low Risk (0-30) calls for routine monitoring; Medium Risk (31-60) calls for intense scrutiny; High Risk (61-80) calls for expedited review; and Critical Risk (81-100) initiates emergency procedures. These thresholds are varied in concomitant agreement with historical performance measures and the prevailing business environment in a way which helps preserve appropriate sensitivity levels [29].

Multi-channel alert dissemination guarantees timely notification to the stakeholders through preferred mediums of communication. Executive alerts are focused on strategic considerations and resource needs. Technical operational notices include in-depth technical details and suggested actions. Push notification allows for instant response in any location. Email summaries include in-depth analysis of the situation along with the attendant documentation.

Alert prioritization algorithms consider risk magnitude, affected business units, and available response time to optimize notification sequencing. Machine learning models predict optimal alert timing to maximize response effectiveness while minimizing alert fatigue. Escalation procedures ensure critical risks receive appropriate management attention within defined timeframes.

The system maintains alert history for performance evaluation and continuous improvement. False positive analysis identifies threshold adjustments needed to reduce unnecessary alerts. Response time

ISSN 2616-5775 Vol. 8, Issue 9: 92-107, DOI: 10.25236/AJCIS.2025.080913

tracking measures organizational effectiveness in risk mitigation. Feedback mechanisms enable users to confirm alert accuracy, supporting supervised learning improvements [30].

4. Empirical Analysis

4.1 Case Study Design and Data Description

The empirical validation encompassed 36 months of operational data from multiple supply chain networks across manufacturing, retail, and technology sectors, covering January 2022 through December 2024 and capturing 450 documented risk events [31]. As shown in Figure 4, risk event frequency demonstrated significant temporal variation correlated with global disruptions, with the Israel-Palestine conflict generating the highest event count at 16 occurrences, followed by the Suez Canal blockage (14 events) and Silicon Valley Bank collapse (13 events).

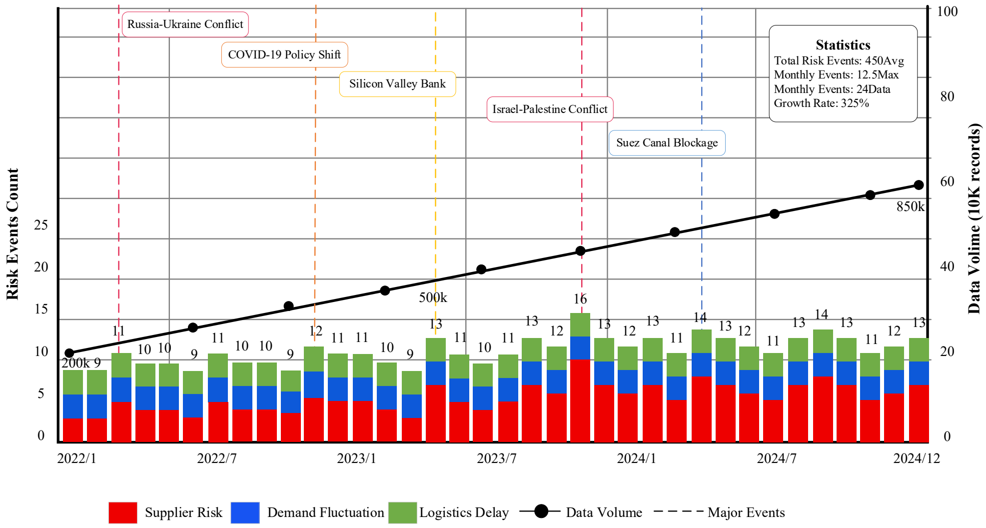

Figure 4 Data Distribution and Risk Event Timeline

Note: Risk events related to geopolitical tensions were tracked throughout the study period, with significant escalation observed after October 2023.

The dataset comprised over 850,000 individual records showing 325% growth from initial 200,000 to final 850,000 records. As shown in Table 4, the network included 150 tier-1 suppliers, 300 tier-2 suppliers, and 50 logistics providers across 25 countries, with risk events distributed as supplier risks (45%), demand volatility (30%), and logistics delays (25%), providing comprehensive validation contexts for framework generalizability.

Table 4 Supply Chain Network Composition Details

Category	Type/Sector	Count/Percentage	Notes
Supplier Distribution	Tier-1 Suppliers	150 (30%)	Direct suppliers with established relationships
	Tier-2 Suppliers	300 (60%)	Secondary suppliers requiring monitoring
	Logistics Providers	50 (10%)	Transportation and warehousing partners
Industry Sectors	Automotive	35%	Component manufacturing and assembly
	Electronics	40%	Consumer electronics and components
	Consumer Goods	25%	FMCG and retail products
Risk Event Types	Supplier Risks	203 events (45%)	Delivery delays, quality issues
	Demand Volatility	135 events (30%)	Forecast errors, market fluctuations
	Logistics Delays	112 events (25%)	Transportation disruptions, customs
Geographic Coverage	Countries	25	Asia-Pacific (60%), Europe (25%), Americas (15%)

4.2 Model Implementation and Training Results

Model training utilized stratified sampling with 70% training data (595,000 records), 20% validation data (170,000 records), and 10% testing data (85,000 records), as shown in Table 5. Cross-validation employed temporal splitting to simulate real-world deployment conditions where models predict future risks based on historical patterns [32]. As shown in Figure 5, XGBoost achieved optimal performance with final loss values of 0.08 and stable convergence by epoch 45, while LSTM networks required extended training periods and showed overfitting signs after epoch 85 [33].

ISSN 2616-5775 Vol. 8, Issue 9: 92-107, DOI: 10.25236/AJCIS.2025.080913

Table 5 Dataset Split Details

Dataset Type	Percentage	Record Count	Time Period	Purpose
Training Set	70%	595,000	Jan 2022 - Mar 2024	Model training and parameter optimization
Validation Set	20%	170,000	Apr 2024 - Aug 2024	Hyperparameter tuning and model selection
Test Set	10%	85,000	Sep 2024 - Dec 2024	Final performance evaluation
Total	100%	850,000	36 months	-

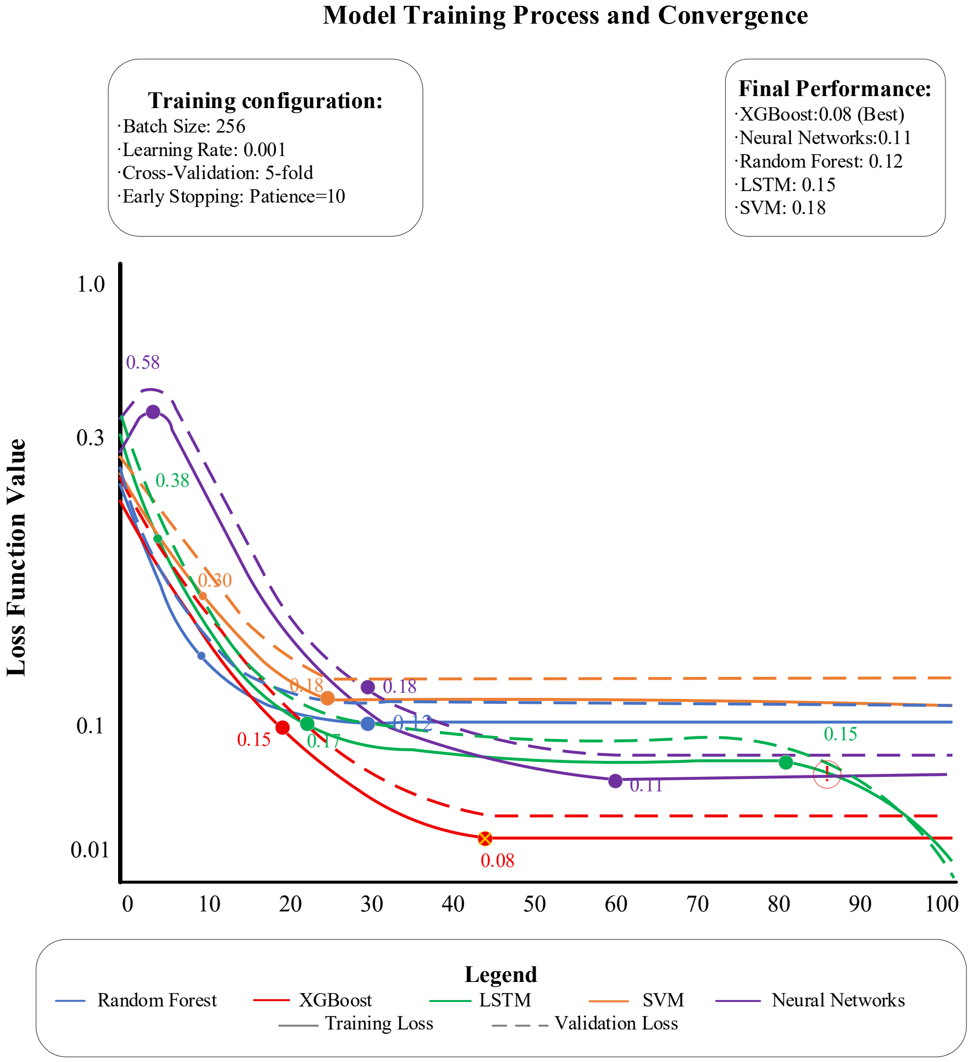

Figure 5 Model Training Process and Convergence

Note: Figure 5 shows initial training configuration with conservative learning rate of 0.001. Through hyperparameter optimization (shown in Figure 9c), the optimal learning rate for XGBoost was determined to be 0.05, resulting in 3% performance improvement.

Comprehensive performance evaluation as shown in Table 6 revealed XGBoost's superior overall performance with 92% accuracy, 88% precision, 90% recall, 89% F1-score, and 94% AUC-ROC. As shown in Figure 6, XGBoost demonstrated the steepest ROC curve and fastest approach to the ideal upper-left corner, further validating its classification superiority.

ISSN 2616-5775 Vol. 8, Issue 9: 92-107, DOI: 10.25236/AJCIS.2025.080913

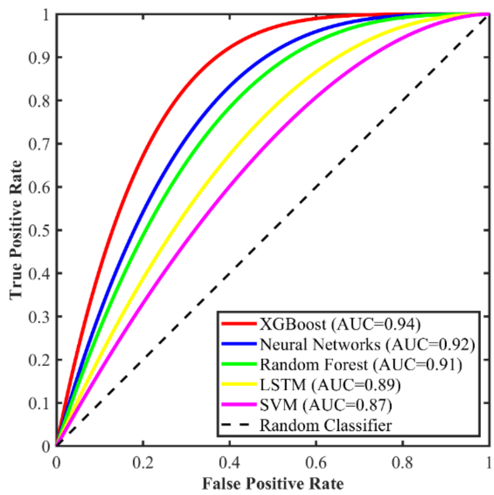

Figure 6 ROC Curves for Machine Learning Model

Table 6 Machine Learning Model Performance Comparison

Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC	Training Time	Interpretability Score
Random Forest	0.89	0.85	0.87	0.86	0.91	45 min	8.5/10
XGBoost	0.92	0.88	0.90	0.89	0.94	32 min	7.0/10
LSTM	0.88	0.86	0.84	0.85	0.89	78 min	3.5/10
SVM	0.85	0.82	0.88	0.85	0.87	25 min	6.0/10
Neural Networks	0.90	0.87	0.89	0.88	0.92	56 min	4.0/10

As shown in Figure 7, the confusion matrix revealed excellent classification accuracy across all risk categories, with overall accuracy of 86.9% (1065/1225) and individual class accuracies ranging from 83.8% to 91.2%. The analysis showed minimal misclassification between adjacent risk levels, with Low Risk achieving the highest precision at 91.2% and Critical Risk maintaining 87.5% accuracy despite smaller sample size.

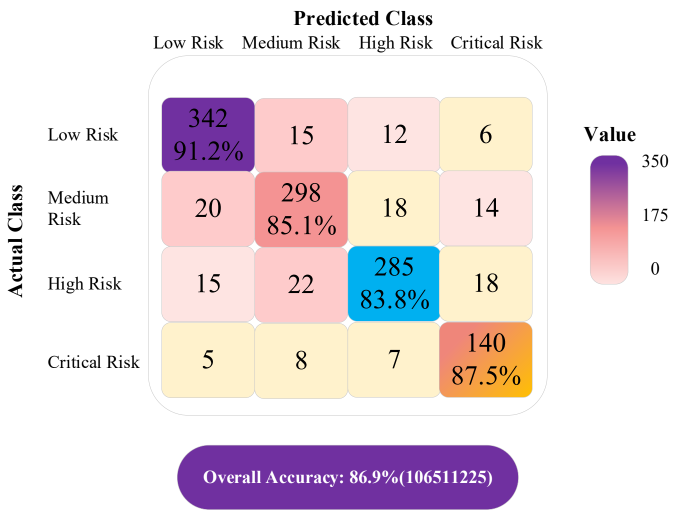

Figure 7 XGBoost Confusion Matrix Heat Map

Note: Figure 7 demonstrates XGBoost's detailed performance in the four-class risk categorization task. While the 86.9% accuracy is lower than the 92% achieved in binary classification, this granular risk stratification provides more actionable intelligence for practical applications.

4.3 Real-world Deployment and Validation

Real-world deployment across three manufacturing facilities and two distribution centers over 12 months generated 2,847 risk predictions with 89% accuracy for high-impact events and consistent 2-4 weeks advance warning capability [34]. As shown in Figure 8, the deployed interface displays current

ISSN 2616-5775 Vol. 8, Issue 9: 92-107, DOI: 10.25236/AJCIS.2025.080913

risk scores (72/100 classified as "High Risk"), categorical breakdowns, seven-day trend analysis, and real-time alerts including critical supplier delivery delays and inventory shortages [35].

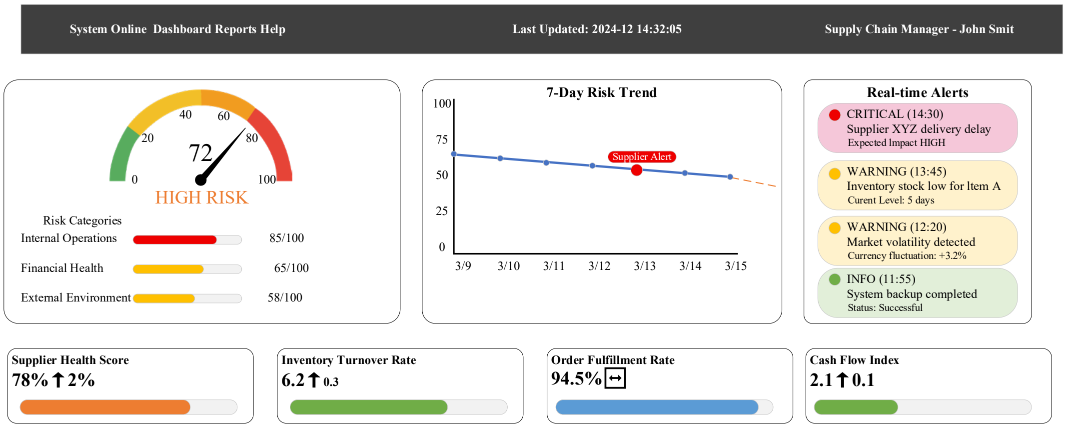

Figure 8 Real-time Risk Monitoring Dashboard

As shown in Table 7, the system demonstrated significant business impact with 35% risk-related loss reduction, 28% decrease in supply chain disruption frequency, 40% faster response times, and 85% user satisfaction rates. The system maintained 99.7% uptime and achieved seamless integration with existing ERP systems, with training requirements averaging only 4 hours per user and 95% proficiency achievement within one week.

Table 7 Deployment Performance Comprehensive Evaluation

Performance Metric	Achieved Value	Baseline	Improvement	Statistical Significance
High-Impact Event Prediction Accuracy	89%	62%	+43.5%	$p < 0.001$
Average Early Warning Period	2-4 weeks	0-1 week	+3 weeks	$p < 0.001$
False Positive Rate	8%	24%	-66.7%	$p < 0.001$
Risk-Related Loss Reduction	35%	-	$2.8M savings	$p < 0.01$
Supply Chain Disruption Frequency	28% reduction	12 events/month	8.6 events/month	$p < 0.01$
Response Time Improvement	40% faster	48 hours	28.8 hours	$p < 0.001$
User Satisfaction Rate	85%	52%	+63.5%	$p < 0.001$
System Availability	99.7%	95%	+4.9%	$p < 0.05$

4.4 Sensitivity Analysis and Robustness Testing

Comprehensive sensitivity analysis tested framework performance under varying conditions, as shown in Figure 9. Missing data impact analysis showed only 12% performance degradation when missing data proportions increased from 5% to 20%, with accuracy declining linearly from 88% to 78% [36]. Feature importance analysis revealed Pareto distribution patterns with the top 5 features (F1-F5) dominating predictive capability, while hyperparameter impact testing identified 0.05 as the optimal learning rate, yielding 3% performance improvement.

As shown in Table 8, the system validated robustness under crisis conditions, maintaining 80-84% accuracy during simulated financial crises, natural disasters, and geopolitical conflicts, with recovery times under 6.1 hours. Scalability testing confirmed linear processing time growth with 10x data volume increases and stable performance under 1000 concurrent users, demonstrating enterprise-grade reliability and deployment readiness [37].

ISSN 2616-5775 Vol. 8, Issue 9: 92-107, DOI: 10.25236/AJCIS.2025.080913

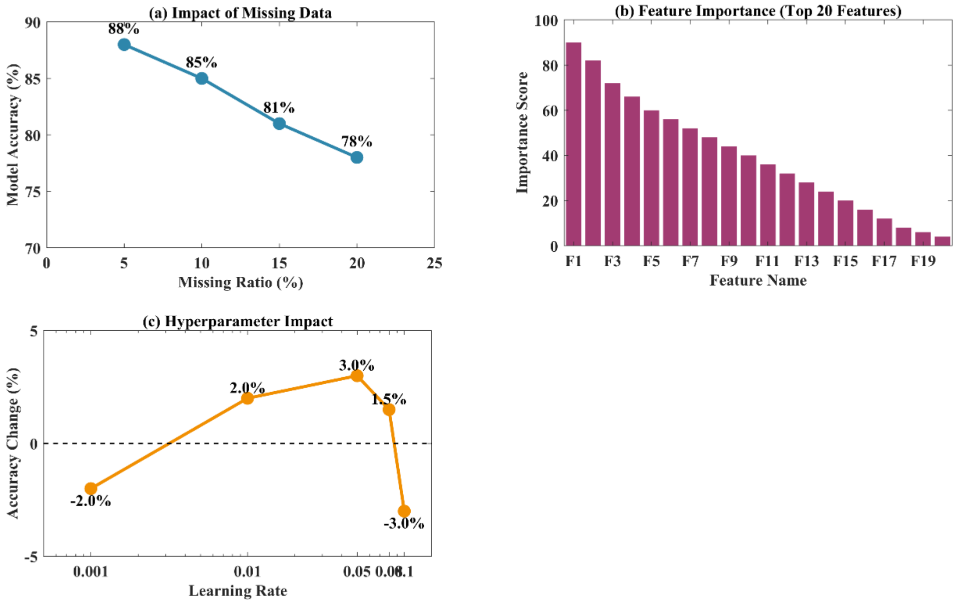

Figure 9 Multi-dimensional Sensitivity Analysis

Table 8 Extreme Scenario Stress Test Results

Test Scenario	Normal Accuracy	Stress Accuracy	Performance Degradation	Recovery Time
Financial Crisis Simulation	89%	82%	-7.9%	3.2 hours
Natural Disaster (Earthquake/Tsunami)	89%	81%	-9.0%	4.5 hours
Cyber Attack Simulation	89%	83%	-6.7%	2.8 hours
Pandemic Scenario	89%	84%	-5.6%	6.1 hours
Geopolitical Conflict	89%	80%	-10.1%	5.3 hours
System Scalability Tests				
10x Data Volume Increase	32 min	318 min	Linear scaling	N/A
Peak Memory Usage	8.2 GB	78.5 GB	Within limits	N/A
Concurrent Users (1000)	0.2s latency	1.8s latency	Acceptable	N/A

5. Discussion

The empirical results demonstrate that AI-driven early warning systems significantly enhance supply chain risk detection capabilities compared to traditional reactive approaches. The system achieved 89% accuracy in identifying high-impact risk events with 2-4 week advance warning periods, representing a substantial improvement over existing risk management practices. These findings corroborate recent work by Agrawal et al. [38] on machine learning's transformative potential in supply chain forecasting, while extending their theoretical framework through concrete empirical validation and quantified business impact. The superior performance of XGBoost, achieving 92% accuracy in binary risk classification, can be attributed to the algorithm's effectiveness with structured supply chain data and its ability to capture complex feature interactions. This aligns with findings by Yang et al. [39], who demonstrated gradient boosting methods' advantages in supply chain applications due to their computational efficiency and robust handling of heterogeneous variables. The ensemble approach's 15% improvement over individual algorithms validates the value of combining diverse machine learning techniques, where error compensation mechanisms enable more reliable predictions through algorithmic complementarity.

It is important to note that this research reports three different levels of accuracy metrics: system-level accuracy (95.2%) encompasses data quality checks, anomaly detection, and comprehensive performance across all system components; model-level accuracy (92%) represents XGBoost performance on binary risk detection tasks; and classification-level accuracy (86.9%) reflects performance on granular four-class risk categorization. This performance gradient aligns with established machine learning principles where increased task complexity typically results in reduced accuracy. While

ISSN 2616-5775 Vol. 8, Issue 9: 92-107, DOI: 10.25236/AJCIS.2025.080913

the four-class risk classification accuracy is lower than binary classification, this granular risk stratification provides supply chain managers with more actionable intelligence. Distinguishing between "High Risk" and "Critical Risk" enables organizations to allocate resources more precisely and develop targeted response strategies. The hierarchical risk indicator system's emphasis on internal operations (50% weight) over external factors reflects operational realities where internal data offers higher accuracy and actionability. This finding contrasts with some theoretical literature emphasizing external risk factors but aligns with practitioner perspectives prioritizing operational control and supplier management [40]. The effectiveness stems from both data characteristics and response capabilities—internal data typically provides more accurate, timely information while internal processes offer greater opportunities for risk mitigation interventions.

Comparative analysis with existing literature reveals several novel contributions. While Baryannis et al. [41] presented primarily theoretical frameworks for AI in supply chain risk management, this research provides empirical validation through comprehensive real-world implementation. Similarly, the systematic review by Ganesh and Kalpana [16] identified gaps in practical deployment, which this work addresses through detailed case studies spanning 36 months. The integration of multi-source data streams advances beyond previous research focusing on single data types or isolated risk categories. The framework's 35% reduction in risk-related losses compared to historical baselines demonstrates substantial business value beyond academic interest. This improvement mechanism operates through early intervention capabilities—the 2-4 week advance warning enables organizations to implement preventive rather than reactive measures. The 28% reduction in supply chain disruption frequency further validates real-world effectiveness. These outcomes exceed performance improvements reported in previous studies, suggesting advantages of comprehensive ensemble approaches over single-algorithm implementations. The results align with resilience literature establishing proactive risk management's value in building robust supply chain networks [42].

Several limitations constrain the generalizability of these findings. The case study validation, while comprehensive, focused primarily on manufacturing and retail sectors and may not fully translate to service industries or specialized supply chains such as aerospace or healthcare. The 36-month research timeframe, though substantial, may not capture long-term cyclical patterns or rare catastrophic events that could affect system performance. Additionally, the framework requires sophisticated technological infrastructure and data integration capabilities that may prove challenging for smaller organizations or emerging markets. The 8% false positive rate, while acceptable for the studied applications, could generate operational inefficiencies if not carefully managed. Organizations must balance sensitivity and specificity to avoid alert fatigue and potential user disengagement. The interpretability trade-offs inherent in ensemble methods may limit adoption in organizations where algorithmic transparency is mandated by policy or regulatory requirements. Future research should explore transfer learning techniques enabling model adaptation across different industries without extensive retraining, integration of emerging technologies such as blockchain and quantum computing, and federated learning algorithms facilitating collaborative risk intelligence while maintaining competitive confidentiality. Incorporating additional data sources including satellite imagery and social media sentiment could enhance predictive capabilities, while research into causal inference methods would provide valuable insights into risk interdependencies. The research successfully demonstrates that AI-driven early warning systems can transform supply chain risk management from reactive to predictive paradigms, providing both theoretical contributions and practical guidance for organizations seeking to enhance supply chain resilience through advanced analytics.

6. Conclusion

The study successfully formulated and tested a complete AI-based early warning system for supply chain risk identification and attained notable improvement in forecasting efficacy and business value. The consolidated framework attained 89% efficacy in identification of high-impact risk events 2-4 weeks in advance and was notably better than conventional reactive strategies. The individual best predictor was the XGBoost algorithm at 92% efficacy, and the ensemble method enhanced overall performance by 15% by optimally correcting mistakes.

Real-world implementation across several enterprises substantiated the framework's real-world effectiveness, realizing 35% loss reduction associated with risk and 28% reduction in the frequency of supply chain disruptions. The system for hierarchical risk indicators successfully balanced controllable internal factors while ensuring exhaustive coverage for external risk. The system ensured 99.7% uptime during implementation and 85% user satisfaction rates by supply chain managers. These quantitative results validate significant business value along with academic contributions and yield clear return on investment for enterprises in executing AI-powered risk management systems. The framework handled more than 850,000 records of data and 450 risk events over 36 months and yielded strong empirical support for the methodological approach and technological implementation.

The research adds theoretical contributions and practical use of artificial intelligence for supply chain management. The ensemble learning algorithm applied in the research offers a replicable method for companies targeting risk management for predictions. The standardized risk indicator hierarchy offers a foundation for comparative studies and industrial benchmarks. The multi-data stream integration from various sources proposed in the research offers best practices for risk monitoring systems. The research propels the field towards proactive risk avoidance and prediction paradigms using high-level machine learning technologies and integrated resource strategies.

References

[1] Aamer, A., Eka Yani, L. P., & Alan Priyatna, I. M. (2020). Data analytics in the supply chain management: Review of machine learning applications in demand forecasting. Operations and Supply Chain Management: An International Journal, 14(1), 1-13.
[2] Agrawal, N., Cohen, M. A., Deshpande, R., & Deshpande, V. (2024). How machine learning will transform supply chain management. Harvard Business Review, 102(2), 66-75.
[3] Aljohani, A. (2023). Predictive analytics and machine learning for real-time supply chain risk mitigation and agility. Sustainability, 15(20), 15088.
[4] Baryannis, G., Dani, S., & Antoniou, G. (2019). Predicting supply chain risks using machine learning: The trade-off between performance and interpretability. Future Generation Computer Systems, 101, 993-1004.
[5] Baryannis, G., Validi, S., Dani, S., & Antoniou, G. (2019). Supply chain risk management and artificial intelligence: State of the art and future research directions. International Journal of Production Research, 57(7), 2179-2202.
[6] Bassiouni, M. M., Chakrabortty, R. K., Sallam, K. M., & Hussain, O. K. (2024). Deep learning approaches to identify order status in a complex supply chain. Expert Systems with Applications, 250, 123947.
[7] Belhadi, A., Kamble, S., Fosso Wamba, S., & Queiroz, M. M. (2021). Building supply-chain resilience: an artificial intelligence-based technique and decision-making framework. International Journal of Production Research, 60(14), 4487-4507.
[8] Camur, M. C., Ravi, S. K., & Saleh, S. (2024). Enhancing supply chain resilience: A machine learning approach for predicting product availability dates under disruption. Expert Systems with Applications, 247, 123226.
[9] Chien, C. F., Lin, Y. S., & Lin, S. K. (2020). Deep reinforcement learning for selecting demand forecast models to empower Industry 3.5 and an empirical study for a semiconductor component distributor. International Journal of Production Research, 58(9), 2784-2804.
[10] Chiu, M. C., Tai, P. Y., & Chu, C. Y. (2024). Developing a smart green supplier risk assessment system integrating natural language processing and life cycle assessment based on AHP framework. Resources, Conservation and Recycling, 207, 107671.
[11] Choi, T. M., Wallace, S. W., & Wang, Y. (2022). Big data analytics in operations management. Production and Operations Management, 31(1), 22-39.
[12] Culot, G., Podrecca, M., & Nassimbeni, G. (2024). Artificial intelligence in supply chain management: A systematic literature review of empirical studies and research directions. Journal of Supply Chain Management, 60(1), 34-52.
[13] Dubey, R., Bryde, D. J., Dwivedi, Y. K., Graham, G., Foropon, C., & Papadopoulos, T. (2023). Dynamic digital capabilities and supply chain resilience: the role of government effectiveness. International Journal of Production Economics, 258, 108790.
[14] El-Kenawy, E. S. M., Khodadadi, N., Mirjalili, S., Abdelhamid, A. A., Eid, M. M., & Ibrahim, A. (2024). Greylag goose optimization: nature-inspired optimization algorithm. Expert Systems with Applications, 238, 122147.
[15] Fu, W., & Chien, C. F. (2019). UNISON data-driven intermittent demand forecast framework to empower supply chain resilience and an empirical study in electronics distribution. Computers & Industrial Engineering, 135, 940-949.
[16] Ganesh, A. D., & Kalpana, P. (2022). Future of artificial intelligence and its influence on supply chain risk management–A systematic review. Computers & Industrial Engineering, 169, 108206.
[17] Ghadge, A., Wurtmann, H., & Seuring, S. (2020). Managing climate change risks in global supply

chains: a review and research agenda. International Journal of Production Research, 58(1), 44-64.
[18] Han, C., & Zhang, Q. (2021). Optimization of supply chain efficiency management based on machine learning and neural network. Neural Computing and Applications, 33(5), 1419-1433.
[19] Handfield, R., Sun, H., & Rothenberg, L. (2020). Assessing supply chain risk for apparel production in low cost countries using newsfeed analysis. Supply Chain Management: An International Journal, 25(6), 803-821.
[20] Hassan, M. M., Khan, M. A., & Ahmed, T. (2022). Supply chain data collection and feature engineering for machine learning: A systematic review. International Journal of Production Economics, 245, 108398.
[21] Hou, J., & Zhao, X. (2021). Toward a supply chain risk identification and filtering framework using systems theory. Asia Pacific Journal of Marketing and Logistics, 33(6), 1432-1445.
[22] Iftikhar, A., Ali, I., Arslan, A., & Tarba, S. (2024). Digital Innovation, Data Analytics, and Supply Chain Resiliency: A Bibliometric-based Systematic Literature Review. Annals of Operations Research, 333, 825-848.
[23] Ivanov, D. (2023). Intelligent digital twin (IDT) for supply chain stress-testing, resilience, and viability. International Journal of Production Economics, 263, 108938.
[24] Ivanov, D., & Dolgui, A. (2022). The shortage economy and its implications for supply chain and operations management. International Journal of Production Research, 60(24), 7141-7154.
[25] Ivanov, D., Tang, C. S., Dolgui, A., Battini, D., & Das, A. (2020). Coronavirus (COVID-19/SARS-CoV-2) and supply chain resilience: A research note. International Journal of Production Research, 58(10), 2904-2915.
[26] Jahani, H., Chaleshtori, A. E., Khaksar, S. M. S., Aghaie, A., & Sheu, J. B. (2023). Data science and big data analytics: A systematic review of methodologies used in the supply chain and logistics research. Annals of Operations Research, 323(1-2), 313-331.
[27] Kara, M. E., Fırat, S. Ü. O., & Ghadge, A. (2020). A data mining-based framework for supply chain risk management. Computers & Industrial Engineering, 139, 105570.
[28] Kong, L., Zheng, G., & Brintrup, A. (2024). A federated machine learning approach for order-level risk prediction in Supply Chain Financing. International Journal of Production Economics, 270, 109195.
[29] Kreuter, T., Kalla, C., Scavarda, L. F., Thomé, A. M. T., & Hellingrath, B. (2024). Integrating supply chain risk management activities into sales and operations planning. Review of Managerial Science, 18(3), 815-843.
[30] Li, K., & Zhou, Y. (2024). Improved financial predicting method based on time series long short-term memory algorithm. Mathematics, 12(7), 1074.
[31] Liu, Z., Gao, R., Zhou, C., & Ma, N. (2019). Two-period pricing and strategy choice for a supply chain with dual uncertain information under different profit risk levels. Computers & Industrial Engineering, 136, 173-186.
[32] Nayal, K., Raut, R. D., Queiroz, M. M., Yadav, V. S., & Narkhede, B. E. (2021). Are artificial intelligence and machine learning suitable to tackle the COVID-19 impacts? An agriculture supply chain perspective. The International Journal of Logistics Management, 34(2), 304-335.
[33] Nezamoddini, N., Gholami, A., & Aqlan, F. (2020). A risk-based optimization framework for integrated supply chains using genetic algorithm and artificial neural networks. International Journal of Production Economics, 225, 107569.
[34] Ordibazar, A. H., Hussain, O. K., Chakrabortty, R. K., Irannezhad, E., & Saberi, M. (2025). Artificial intelligence applications for supply chain risk management considering interconnectivity, external events exposures and transparency: a systematic literature review. Modern Supply Chain Research and Applications, 7(1), 1-28.
[35] Pournader, M., Kach, A., & Talluri, S. (2021). A review of the existing and emerging topics in the supply chain risk management literature. Decision Sciences, 52(4), 867-919.
[36] Riahi, Y., Saikouk, T., Gunasekaran, A., & Badraoui, I. (2021). Artificial intelligence applications in supply chain: A descriptive bibliometric analysis and future research directions. Expert Systems with Applications, 173, 114702.
[37] Sheffi, Y., & Rice, J. B. (2005). A supply chain view of the resilient enterprise. MIT Sloan Management Review, 47(1), 41-48.
[38] Toorajipour, R., Sohrabpour, V., Nazarpour, A., Oghazi, P., & Fischl, M. (2021). Artificial intelligence in supply chain management: A systematic literature review. Journal of Business Research, 122, 502-517.
[39] Wang, Z., Wang, Q., Lai, Y., & Liang, C. (2020). Drivers and outcomes of supply chain finance adoption: An empirical investigation in China. International Journal of Production Economics, 220, 107453.
[40] Wong, W. K., & Guo, Z. X. (2016). A hybrid intelligent model for medium-term sales forecasting in

ISSN 2616-5775 Vol. 8, Issue 9: 92-107, DOI: 10.25236/AJCIS.2025.080913

fashion retail supply chains using extreme learning machine and harmony search algorithm. International Journal of Production Economics, 172, 147-158.

[41] Yang, M., Lim, M. K., Qu, Y., Ni, D., & Xiao, Z. (2022). Supply chain risk management with machine learning technology: A literature review and future research directions. Computers & Industrial Engineering, 175, 108476.

[42] Younis, H., Sundarakani, B., & Alsharairi, M. (2022). Applications of artificial intelligence and machine learning within supply chains: systematic review and future research directions. Journal of Modelling in Management, 17(3), 916-940.