Information Security Education Journal (ISEJ)

DLINE JOURNALS

Print ISSN: 2349-8161
Online ISSN: 2349-817X

ISEJ 2026:13 (2)
https://doi.org/10.6025/isej/2026/13/2/122-147

Machine-Learning-Based Prediction of Production Disruptions in Supply Chains

Hajar Ait Lamkademe
University of Hassan II of Casablanca. Morocco
hajar.aitlamkademe-etu@etu.univh2c.ma

ABSTRACT

Global supply chains face increasing exposure to cybersecurity related disruptions that can interrupt production and degrade operational continuity. Conventional risk assessment approaches often rely on lagging indicators and expert judgement, limiting their ability to provide timely, quantifiable estimates of production impact. This study develops and evaluates an explainable machine learning framework for estimating the percentage of production affected by cybersecurity incidents. Using a dataset of 85,000 observations from 4,000 suppliers spanning multiple industries and regions, eight regression algorithms (Linear Regression, Ridge, Lasso, Support Vector Regression, Random Forest, Gradient Boosting, XG Boost, and Artificial Neural Network) are trained and compared under a rigorous supplier grouped validation protocol that ensures complete separation of suppliers between training and test sets.

Gradient Boosting Regressor achieved the strongest predictive performance (MAE = 9.21, RMSE = 13.35, R^2= 0.768), with Random Forest and XGBoost performing nearly as well. Feature importance, SHAP, and partial-dependence analyses identified network and cybersecurity indicators particularly missed_bytes—as the dominant predictors, while sensitivity experiments showed that meaningful predictive power is retained even after the removal of this variable. The results demonstrate that supplier, network traffic, and cybersecurity features contain substantial signal for estimating production impact on previously unseen suppliers. At the same time, the non chronological validation design and the measurement timing of key predictors constrain claims of prospective early warning capability. The study contributes a transparent, supplier independent modelling pipeline and highlights the methodological requirements for advancing from retrospective impact estimation toward operational early warning systems.

Keywords: Supply-Chain Disruption, Production-Impact Prediction, Cybersecurity Risk, Machine Learning, Gradient Boosting, Explainable AI, SHAP, Supplier-Grouped Validation, Network Traffic Indicators

Received: 19 March 2026, Revised: 15 August 2026, Accepted: 29 August 2026

Copyright: DLINE

1. Introduction

Global supply chains are increasingly exposed to disruptions arising from geopolitical conflicts, fluctuations in demand, transportation failures, and environmental instability. These interacting pressures create vulnerabilities across interconnected logistics and manufacturing networks and can undermine the continuity of production and the reliability of supply. Conventional risk-assessment practices, which often depend on manual evaluation and historical observations, may be insufficient for identifying rapidly evolving threats and providing timely warnings. In this context, machine learning (ML) offers opportunities to strengthen predictive intelligence and support more informed decision-making in complex supply-chain environments [1].

2. Background

2.1 Machine Learning for Early Warning and Supply-Chain Risk Assessment

Recent work has explored the use of ML-based early warning systems to identify emerging operational problems. One study describes an approach for detecting and predicting defects in semiconductor devices and situates this capability within a broader framework for supply chain risk assessment and mitigation. The proposed system draws on multiple data streams, including real time measurement data, information from performance-measurement equipment, and business indicators. It also refers to the combined use of graph based network modelling and sequence learning methods, specifically Graph Neural Networks (GNNs) and Long Short Term Memory (LSTM) networks, to represent network relationships and temporal patterns [2].

Supplier related risks constitute another important dimension of supply chain vulnerability. The growing complexity and global interdependence of manufacturing networks have increased exposure to supplier disruptions associated with financial instability, operational failures, geopolitical tensions, and natural disasters. Static, rule-based supplier scorecards and expert judgement remain common evaluation approaches; however, these methods may not adequately represent changes in risk over time or the resilience behaviour of interconnected suppliers. This limitation motivates more dynamic approaches to supplier risk evaluation [3].

A related challenge concerns the delayed nature of many conventional monitoring indicators. Transportation congestion, logistics imbalances, and broader macroeconomic shocks can spread through freight networks, while monitoring systems based primarily on lagging metrics may fail to provide actionable early warning signals. To address this issue, Sarker describes a multimodal prediction and risk assessment pipeline that combines the Global Supply Chain Pressure Index (GSCPI), freight indicators from the Bureau of Transportation Statistics, transportation time series from the Federal Reserve Economic Data (FRED) system, and derived anomaly signals. The reported benchmarking includes temporal and non-temporal approaches, such as LSTM forecasting, Extreme Gradient Boosting (XGBoost), Random Forest, and an Isolation Forest based anomaly detection layer. The pipeline is intended to identify emerging stress conditions and predict disruption states over short forecasting horizons [4].

2.2 Disruption Propagation and Supply-Chain Resilience Modelling

In addition to predicting disruption events, prior research has examined how disruptions propagate through supply networks and how their consequences can be mitigated through resilience oriented optimisation and simulation. Linear programming has been used to analyse domino effects in supply chains, providing a means of examining how an initial disruption may affect connected operations [5]. Discrete event simulation has also been applied to investigate ripple effects while accounting for disruption duration, supplier delays, and the resulting reduction in service levels [6, 7]. Recovery time is another relevant factor in modelling disruption consequences and the restoration of supply chain operations [8].

Other resilience studies extend ripple-effect analysis by incorporating recovery costs and their influence on service levels. They also consider the structure of supplier networks as a means of understanding or preventing disruption propagation [9, 10]. Inventory and transportation costs have further been incorporated into analyses

of delivery delays and backorders, including studies based on discrete event simulation. Probabilistic approaches, such as Bayesian networks and Markov chains, have likewise been used to evaluate supply chain service levels under uncertain operating conditions [11, 12]. Supplier disruption risk has also been examined through decision support systems and Bayesian hierarchical modelling [13, 14, 15, 16],

2.3 Synthesis and Research Motivation

The reviewed studies indicate two complementary strands of supply chain disruption research. The first focuses on predictive monitoring and early warning, using heterogeneous operational, transportation, business, and macro level data together with ML methods to detect emerging stress and estimate near term disruption states. The second focuses on disruption propagation and resilience, using optimisation, simulation, and probabilistic models to examine domino and ripple effects, service level degradation, recovery requirements, and associated costs. Together, these strands establish the importance of both anticipating disruption and understanding its consequences across interconnected supply chain operations.

A research opportunity follows from the need to connect predictive disruption intelligence with the operational consequences of disruption propagation and recovery. The material reviewed here describes ML-based early-warning approaches as well as resilience models, but does not itself specify a unified predictive and resilience framework, a common evaluation dataset, or a comparative experimental protocol. A study addressing production disruption prediction could therefore investigate how multi source operational and network information can be integrated to identify disruption risk early, while relating predicted risk to production continuity, supplier dependencies, service levels, and recovery behaviour. This motivation is drawn from the themes in the supplied material; the precise research questions, variables, model architecture, and evaluation design would need to be defined based on the study's dataset and scope.

2.4 Research Statement

2.4.1 Research Problem

The increasing dependence of manufacturing supply chains on interconnected digital infrastructure has created additional vulnerabilities to cybersecurity related disruptions. Cybersecurity incidents can affect supplier operations, network communication, information availability, and production continuity, resulting in varying degrees of operational losses. The consequences are not uniform: some incidents may produce no measurable production impact, whereas others may substantially affect manufacturing activities and supply chain performance.

Traditional supply chain risk assessment approaches frequently depend on predefined indicators, historical records, and expert based evaluations. Although these approaches provide useful information, they may be insufficient to capture complex relationships between supplier characteristics, network traffic behaviour, cybersecurity indicators, and production consequences.

The central research problem addressed in this study is therefore the development of a machine learning based predictive framework capable of estimating the percentage of production affected by cybersecurity incidents. Specifically, the study investigates whether supplier related characteristics, network traffic measurements, cybersecurity indicators, and temporal attributes can be used to estimate production impact across different suppliers.

An additional methodological challenge concerns the generalisation of predictive models to previously unseen suppliers. Conventional random train–test partitioning may introduce supplier related information overlap, potentially producing optimistic estimates of model performance. To address this issue, the present study adopts supplier grouped validation, ensuring that suppliers represented in the testing dataset are completely excluded from model training.

Furthermore, the study investigates model interpretability and predictor dependency. Particular attention is given to the variable missed_bytes, which demonstrates exceptionally high predictive importance. Its contribution is examined through feature importance analysis, SHAP interpretation, partial dependence analysis, and sensitivity experiments.

2.5 Research Gap

The literature review identifies the following research gaps:

Gap 1: Limited integration of cybersecurity indicators and production impact estimation. Existing research has investigated machine learning based supply chain risk assessment and early warning systems. However, there remains a need to systematically examine how cybersecurity related network and supplier indicators can estimate the magnitude of production impact.

Gap 2: Insufficient consideration of supplier-independent generalisation. Supply chain datasets contain multiple observations associated with individual suppliers. Random observation level partitioning may allow supplier-specific characteristics to appear in both training and testing datasets. Supplier grouped validation provides a more rigorous assessment of generalisation to unseen suppliers.

Gap 3: Limited interpretability of production-impact prediction. High predictive accuracy alone does not explain which cybersecurity and network indicators influence estimated production losses. Explainable machine-learning techniques are required to investigate predictor contributions and model behaviour.

Gap 4: Potential dependence on dominant network indicators. The contribution of individual network measurements may be disproportionately large. Such dependence requires sensitivity analysis to determine whether the model retains useful predictive capability when dominant indicators are removed.

Gap 5: Distinction between retrospective estimation and prospective early warning. A model may achieve high predictive performance using measurements recorded during or after an incident. However, such performance does not necessarily demonstrate the ability to predict disruption before it occurs. Predictor availability and measurement timing must therefore be considered when assessing operational applicability.

2.6 Research Aim

The primary aim of this research is to develop and experimentally evaluate an explainable machine learning framework for estimating cybersecurity-related production disruption in supply chains by integrating supplier characteristics, network traffic indicators, cybersecurity measurements, and temporal information.

2.7 Research Objectives

The study addresses the following objectives:

1. To investigate the relationships between supplier characteristics, network traffic behaviour, cybersecurity indicators, and production impact.

2. To formulate production-impact estimation as a supervised regression problem using production_impact_pct as the continuous dependent variable.

3. To develop and compare eight machine-learning regression algorithms for estimating production impact.

4. To evaluate model generalisation using supplier-grouped training and testing partitions.

5. To identify influential predictors using feature importance, SHAP values, and partial dependence analysis.

6. To examine prediction errors through residual analysis and observed-versus-predicted comparisons.

7. To investigate the sensitivity of the selected model to the removal of the dominant predictor, missed_bytes.

8. To assess the methodological requirements for extending the framework towards prospective supply-chain disruption early warning.

2.8 Research Questions

The following research questions guide the investigation:

• RQ1: To what extent can supplier, network traffic, and cybersecurity indicators estimate the percentage of production affected by cybersecurity incidents?

• RQ2: Which regression algorithm provides the most effective predictive performance when evaluated on previously unseen suppliers?

• RQ3: Which variables contribute most significantly to the model's production-impact estimates?

• RQ4: How does removing missed_bytes influence predictive accuracy and model robustness?

• RQ5: What prediction patterns and error characteristics emerge across different levels of production impact?

• RQ6: To what extent can the developed framework support supplier independent production-impact assessment, and what additional validation is required for prospective early warning applications?

3. Methodology

3.1 Overall Research Methodology

This research adopts a quantitative, data driven experimental methodology based on supervised machine learning. The proposed approach consists of data preparation, predictor selection, leakage control, preprocessing, model development, supplier grouped validation, performance evaluation, explainability analysis, and sensitivity assessment.

The prediction task is formulated as a regression problem in which the objective is to estimate the continuous percentage of production affected by a cybersecurity incident.

The dataset is represented as

$$D = \{(X_i, y_i)\}_{i=1}^{N}$$

where:

• N is the total number of observations,

• X_i s the predictor vector associated with observation i,

• y_i is the observed production-impact percentage.

The target variable is defined as

y_i = \text{production_impact_pct}_i.

The predictive relationship is expressed as

$$\hat{y}_i = f(x_i;\theta)$$

where f is the selected machine-learning algorithm, θ denotes its learned parameters, and $\hat{y}_i$ is the estimated production impact.

3.2 Dataset Description

The dataset comprises 85,000 observations involving 4,000 unique suppliers, with cybersecurity

incidents recorded between January and June 2024. It contains 32 original variables describing supplier characteristics, network traffic, cybersecurity activities, threat characteristics, operational consequences, and incident response.

The dataset covers eight industries, five geographical regions, three supplier-size categories, and three supply-chain tiers.

The predictor information is organised into the following groups:

1. Supplier characteristics: Industry, geographic region, firm size, supply-chain tier, and backup-supplier availability.

2. Network traffic characteristics: Session duration, packet rate, byte rate, connection state, and missed bytes.

3. Cybersecurity indicators: Checksum errors, OSSEC alerts, login attempts, successful logins, file activity, process activity, privileged-access events, MFA status, device compliance, and patch level compliance.

4. Temporal characteristics: Month, weekday, and hour derived from incident timestamps.

The dependent variable, production_impact_pct, represents the percentage of production affected by an incident.

The target contains 85,000 valid observations, with a mean of 29.42%, standard deviation of 27.59 percentage points, median of 25.20%, and observed range of 0–95%. Approximately 25.1% of the observations have zero recorded production impact.

3.3 Data Preprocessing and Leakage Control

Data preprocessing is conducted to improve consistency and ensure compatibility with the selected regression algorithms.

Numerical missing values are handled using median imputation, while categorical missing values are replaced using the mode. Categorical predictors are transformed through one-hot encoding to enable their inclusion in machine-learning models.

Standardisation is applied to Linear Regression, Ridge Regression, Lasso Regression, Support Vector Regression, and Artificial Neural Network models.

All preprocessing parameters should be estimated using training observations only and subsequently applied to the testing dataset to avoid information leakage.

To prevent target leakage, the following variables are excluded from the predictor set:

- Event and supplier identifiers.
- Attack type and threat label.
- Disruption severity.
- Revenue loss.
- Response strategy.
- Response time.

• Recovery days.

These variables are excluded because they may encode the incident outcome or represent downstream consequences rather than independent predictive information.

The timestamp is transformed into month, weekday, and hour to capture temporal patterns without directly using event identifiers.

The measurement timing of missed_bytes is treated as a separate methodological consideration because its availability during or after an incident could influence whether the model is appropriate for retrospective estimation or advance warning.

3.4 Machine-Learning Model Development

Eight regression algorithms are implemented to compare predictive performance across linear, nonlinear, ensemble, kernel-based, and neural approaches.

3.41 Linear Regression

Linear Regression is used as a baseline model to estimate the relationship between the predictor variables and production impact:

$$\hat{y} = \beta_0 + \sum_{j=1}^{p} \beta_j x_j$$

where β_0 is the intercept and β_j is the coefficient associated with predictor j.

3.42 Ridge Regression

Ridge Regression incorporates L_2 regularisation to constrain coefficient magnitudes and reduce model complexity:

$$L(\beta) = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^{p} \beta_j^2$$

3.43 Lasso Regression

Lasso Regression uses L_1 regularisation to penalise coefficient magnitudes and potentially eliminate less informative predictors:

$$L(\beta) = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^{p} |\beta_j|$$

3.43 Support Vector Regression

An RBF-kernel Support Vector Regression model is employed to represent nonlinear relationships between cybersecurity indicators and production impact.

The SVR model uses a deterministic 5,000-observation subset of the training data because of computational constraints.

5. Random Forest Regression

Random Forest combines multiple decision trees and aggregates their predictions to capture nonlinear relationships and interactions among predictors.

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} f_t(X)$$

where T represents the number of trees and $f_t(X)$ denotes the prediction from tree t .

3.44 Gradient Boosting Regression

Gradient Boosting develops an ensemble of sequential decision trees, with successive learners reducing prediction errors from previous iterations.

Gradient Boosting update rule

$$F_m(X) = F_{m-1}(X) + \eta\, h_m(X)$$

where η is the learning rate and $h_m(X)$ denotes the additional (weak) learner.

3.45 XGBoost Regression

XGBoost is employed as a regularised gradient-boosting algorithm to model nonlinear relationships and interactions among supplier, network, and cybersecurity variables.

3.46 Artificial Neural Network

An Artificial Neural Network is used to learn nonlinear mappings between the input predictor space and continuous production impact values through interconnected computational units and learned weights.

The precise network architecture and hyperparameter configurations should be documented from the original implementation.

3.5 Model Evaluation

The predictive performance of all eight algorithms is evaluated using four regression metrics.

Mean Absolute Error (MAE)

$$\mathbf{MAE} = \frac{1}{N}\sum_{i=1}^{N} |\, y_i - \hat{y}_i \,|$$

MAE measures the average absolute deviation between observed and predicted production impact.

Mean Squared Error (MSE)

$$\mathbf{MSE} = \frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y}_i)^2$$

MSE penalises larger prediction errors more strongly than smaller errors.

Root Mean Squared Error (RMSE)

$$\mathbf{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}$$

RMSE provides an error measure in the original production-impact percentage point scale.

Coefficient of Determination (R^2)

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}$$

A higher (R^2) indicates that the model explains a larger proportion of the observed variation in production impact.

3.6 Explainable Machine Learning

Explainability analysis is conducted to investigate the predictive behaviour of the selected Gradient Boosting model.

Three complementary approaches are incorporated.

Feature Importance: Tree-based feature importance is used to identify the variables contributing most strongly to the fitted model.

SHAP Analysis: SHapley Additive exPlanations are employed to examine the contribution of individual predictors to model predictions. Positive SHAP values increase estimated production impact, whereas negative values reduce it.

Partial Dependence Plots: PDP analysis investigates how the model's average prediction changes as selected predictor values vary.

The explainability analysis focuses particularly on missed_bytes, ossec_alert_level, successful_logins, and paket_rate.

These techniques explain model behaviour and predictive associations rather than establishing causal relationships.

3.7 Sensitivity Analysis

A feature-ablation experiment is conducted to assess the dependence of the selected Gradient Boosting model on missed_bytes.

Two model specifications are evaluated:

- Full model: Includes all selected predictors, including missed_bytes.
- Reduced model: Excludes missed_bytes while retaining the remaining predictors.

Both specifications are evaluated using the same supplier grouped test partition and the same performance metrics.

The resulting differences in MAE, MSE, RMSE, and (R^2) provide an indication of the predictive contribution of the dominant network related feature.

4. Research Design and Proposed Architecture

4.1 Research Design

The study follows a comparative experimental research design based on supervised machine learning regression.

The research design integrates four principal analytical components:

1. Predictive modelling of production impact.

2. Comparative evaluation of eight regression algorithms.

3. Explainability and prediction diagnostics.

4. Sensitivity analysis through dominant feature exclusion.

The design employs supplier-level grouping to ensure that the training and testing datasets contain entirely different suppliers. This provides a more demanding evaluation of model generalisation than conventional observation-level random partitioning.

4.2 Proposed Framework Architecture

The proposed framework consists of the following sequential modules:

Module 1: Supply-Chain Cybersecurity Data Acquisition

Collect incident-level information containing supplier characteristics, network traffic measurements, cybersecurity indicators, and timestamp information.

Module 2: Data Quality Assessment and Preprocessing

Examine missing values, duplicate records, variable types, data distributions, and inconsistencies. Apply numerical imputation, categorical imputation, one-hot encoding, and appropriate standardisation.

Module 3: Leakage-Controlled Feature Engineering

Select supplier, network, cybersecurity, and temporal predictors while excluding identifiers and potentially post-outcome variables.

Module 4: Supplier-Grouped Data Partitioning

Separate the dataset into training and testing groups based on supplier identity, ensuring zero supplier overlap.

Module 5: Parallel Regression Model Development

Train Linear Regression, Ridge Regression, Lasso Regression, SVR, Random Forest, Gradient Boosting, XGBoost, and ANN models.

Module 6: Predictive Performance Evaluation

Evaluate all models using MAE, MSE, RMSE, and (R^2) on the same held-out suppliers.

Module 7: Explainable Prediction Analysis

Apply feature importance, SHAP analysis, and partial dependence plots to investigate the contribution and behaviour of important predictors.

Module 8: Robustness and Sensitivity Assessment

Remove missed_bytes and compare the resulting performance with the full model.

Module 9: Production-Impact Assessment

Interpret the estimated production-impact percentages and examine prediction errors across different impact levels.

4.3 Architecture Representation

The overall workflow can be represented as:

Data Acquisition → Data Quality Assessment → Preprocessing and Feature Engineering →Leakage Control → Supplier Grouped Train Test Partitioning → Eight Regression Models →Performance Evaluation → Explainable

ML → Sensitivity Analysis → Production-Impact Assessment.

The architecture is designed to connect cybersecurity related supplier information with quantitative estimation of operational production consequences.

5. Experimental Setup

5.1 Experimental Dataset

The experimental evaluation uses 85,000 observations associated with 4,000 suppliers. The dataset includes 32 original variables and covers cybersecurity incidents recorded between January and June 2024.

The target variable is production_impact_pct, representing the percentage of production affected.

The dataset contains no reported duplicate rows or duplicate event identifiers, and the target variable contains no missing values.

5.2 Training and Testing Configuration

An 80:20 supplier-grouped partition is implemented using GroupShuffleSplit with a fixed random state of 42.

The resulting experimental configuration is:

Parameter	Configuration
Total observations	85,000
Total suppliers	4,000
Training observations	68,004
Testing observations	16,996
Training suppliers	3,200
Testing suppliers	800
Supplier overlap	0
Training to testing ratio	80:20
Random state	42

The absence of supplier overlap ensures that testing observations originate exclusively from suppliers not represented during training.

This protocol evaluates cross supplier generalisation. It does not establish chronological forecasting performance.

5.3 Experimental Model Configuration

Eight regression algorithms are evaluated under the same predictor framework and common testing

observations.

Numerical predictors are median imputed, categorical predictors are mode imputed and one hot encoded, and standardisation is applied to the linear, SVR, and ANN models.

The RBF-SVR uses a deterministic 5,000-row subset of the training observations because of computational requirements. Other models use the complete training subset.

Exact hyperparameter values and the ANN architecture should be reported from the original experimental implementation.

5.4 Comparative Model Evaluation

The reported experimental results are:

Model	**MAE**	**RMSE**	**R^2**
Gradient Boosting	9.210	13.352	0.7679
Random Forest	9.213	13.365	0.7675
XGBoost	9.262	13.407	0.7660
ANN	9.380	13.505	0.7626
Ridge Regression	10.188	13.786	0.7526
Linear Regression	10.187	13.786	0.7526
Lasso Regression	10.188	13.786	0.7526
SVR	9.742	14.159	0.7391

Gradient Boosting achieved the strongest observed overall performance, with an (R^2) of 0.7679 and an RMSE of 13.352 percentage points.

The closely comparable performance of Random Forest and XGBoost indicates that multiple ensemble-based approaches can capture useful nonlinear predictive relationships in the dataset.

These results represent performance on one held-out supplier partition and should not be interpreted as statistically significant superiority without repeated validation.

5.5 Explainability and Sensitivity Experiments

The selected Gradient Boosting model is subjected to feature-importance analysis, SHAP interpretation, partial dependence analysis, and prediction diagnostics.

The dominant predictor identified in the analysis is missed_bytes, which accounts for approximately 99.8% of the aggregated impurity-based importance.

The sensitivity experiment produces the following results:

Specification	MAE	RMSE	R^2
With missed_bytes	9.210	13.352	0.7679
Without missed_bytes	9.668	14.070	0.7423

Removing missed_bytes reduces the explained variance by approximately 2.56 percentage points. Nevertheless, the reduced model retains considerable predictive performance, suggesting that the remaining supplier, network, and cybersecurity variables also contain useful predictive information.

5.6 Prediction Diagnostics

The selected model is additionally examined using:

- Observed-versus-predicted production-impact plots.
- Residual distribution analysis.
- Residuals-versus-predicted plots.
- SHAP summary plots.
- Partial dependence plots.
- Observed-versus-predicted comparisons across production-impact categories.

These analyses provide complementary evidence concerning prediction accuracy, error distribution, predictor contributions, and differences in predictive behaviour across impact levels.

5.7 Reproducibility Requirements

For reproducibility, the final experimental report should specify the programming environment, software package versions, hardware configuration, complete hyperparameter settings, and ANN architecture using the original implementation records.

The reported dataset partition, random state, preprocessing procedures, evaluation metrics, and SVR subsampling strategy should be retained consistently.

6. Results

6.1 Model comparison

Model	MAE	MSE	RMSE	R^2
Gradient Boosting Regressor	9.210	178.283	13.352	0.7679
Random Forest Regressor	9.213	178.621	13.365	0.7675
XGBoost Regressor	9.262	179.744	13.407	0.7660

Artificial Neural Network	9.380	182.394	13.505	0.7626
Ridge Regression	10.188	190.045	13.786	0.7526
Linear Regression	10.187	190.046	13.786	0.7526
Lasso Regression	10.188	190.046	13.786	0.7526
Support Vector Regression	9.742	200.469	14.159	0.7391

Table 1. Performance comparison of machine-learning regression models

Table 1 presents the predictive performance of eight regression algorithms evaluated for estimating the percentage of production affected by cybersecurity-related supply-chain incidents. The models were assessed using mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and the coefficient of determination (R^2). These metrics provide complementary perspectives on prediction accuracy, the magnitude of prediction errors, and the proportion of variance in production impact explained by each model. The Gradient Boosting Regressor achieved the strongest overall performance, recording an MAE of 9.210, MSE of 178.283, RMSE of 13.352, and R^2 of 0.7679. This indicates that its predictions deviated from the observed production-impact values by approximately 9.21 percentage points on average and explained approximately 76.79% of the variation in the target variable within the held-out test dataset. Random Forest achieved closely comparable results, with an MAE of 9.213 and R^2 of 0.7675. XGBoost also demonstrated competitive performance, achieving an R^2 of 0.7660.

The Artificial Neural Network produced an R^2 of 0.7626, indicating reasonably strong predictive capability, although its performance was marginally lower than that of the leading ensemble-based models. Ridge, Linear, and Lasso Regression produced nearly identical R^2 values of approximately 0.7526. Support Vector Regression recorded the lowest R^2 of 0.7391 and the highest RMSE of 14.159 among the evaluated models.

Overall, the results indicate that ensemble based tree algorithms were particularly effective in capturing the nonlinear relationships and interactions associated with production disruption. Gradient Boosting was selected as the best performing model based on the observed test set metrics. However, the differences between Gradient Boosting, Random Forest, and XGBoost were relatively small and should not be interpreted as evidence of statistically significant superiority without repeated validation and uncertainty estimates.

index	**Value**
count	85000
mean	29.41892353
std	27.58790951
min	0
25%	0
50%	25.2
75%	53.1

90%	68.4
95%	77.8
99%	89.7
max	95

Table 2. Statistical distribution of the production impact target variable

Table 2 summarizes the statistical characteristics of the target variable, production_impact_pct, which represents the percentage of production affected by a cybersecurity incident. The dataset contains 85,000 observations, with a mean production impact of 29.418% and a standard deviation of 27.587 percentage points. The median is 25.2%, indicating that at least half of the recorded incidents affected no more than approximately one-quarter of production.

Table 2. Statistical distribution of the production-impact target variableTable 2 summarizes the statistical characteristics of the target variable, production_impact_pct, which represents the percentage of production affected by a cybersecurity incident. The dataset contains 85,000 observations, with a mean production impact of 29.418% and a standard deviation of 27.587 percentage points. The median is 25.2%, indicating that at least half of the recorded incidents affected no more than approximately one-quarter of production.The distribution extends from 0% to 95%, demonstrating substantial variation in operational consequences across incidents. The first quartile is 0%, while the third quartile is 53.1%, indicating that a considerable proportion of observations experienced no recorded production impact, whereas a substantial upper segment experienced moderate to high disruption. The 90th, 95th, and 99th percentiles are 68.4%, 77.8%, and 89.7%, respectively. These values demonstrate the presence of incidents associated with particularly severe production consequences.

The 25.1% zero-impact proportion is especially important for model development. It indicates that the target distribution includes a meaningful concentration of observations with no recorded production disruption. Consequently, the prediction problem combines the identification of zero-impact events with the estimation of continuous production losses among affected observations. This characteristic may partly explain why individual regression algorithms exhibit different error patterns. It also provides a justification for investigating two-stage modelling approaches that first classify whether production disruption occurs and subsequently estimate its magnitude when the impact is positive.

Item	**Value**
Total observations	85000
Training observations	68004
Test observations	16996
Total suppliers	4000
Training suppliers	3200
Test suppliers	800

Supplier overlap	0
Test target mean	29.64492233
Test target SD	27.71876184
Test target zero proportion	0.249882325

Table 3. Validation Protocol

Table 3. Supplier-grouped training and testing validation protocol

Table 3 documents the dataset partitioning and validation strategy used to evaluate whether the developed models can generalize to suppliers that were not represented during model training. An 80:20 supplier-grouped split was implemented using GroupShuffleSplit, with a fixed random state of 42. The training subset consisted of 68,004 observations from 3,200 suppliers, while the test subset contained 16,996 observations from 800 suppliers.

Importantly, the supplier overlap between the training and testing subsets was zero. This confirms that all suppliers represented in the test dataset were excluded from the training process. Such a strategy is particularly relevant to supply-chain cybersecurity because observations originating from the same supplier may share operational, network, and organizational characteristics. Randomly distributing individual observations across training and testing subsets could therefore result in overly optimistic estimates of predictive performance.

The supplier grouped validation strategy provides a more stringent assessment of the models' ability to predict production disruption for previously unseen suppliers. The test-set target mean, standard deviation, and zero-impact proportion additionally characterize the distribution of the independent evaluation data. The preprocessing procedures, including median imputation for numerical variables, mode imputation and one-hot encoding for categorical variables, and standardization for selected algorithms, were applied as part of the modelling pipeline.

An important qualification is that this protocol evaluates generalization across suppliers rather than prediction of future incidents. Since the partition was not chronological, the results should not be interpreted as evidence of temporal forecasting performance.

Original variable	**Importance**
missed_bytes	0.998093
ossec_alert_level	0.000299
successful_logins	0.000253
paket_rate	0.000223
duration_s	0.000188
byte_rate	0.000179

patch_level	0.000161
file_activity	0.0001
ossec_alerts	0.000088
device_compliance	0.000084
login_attempts	0.00008
process_activity	0.000077
conn_state	0.000043
has_backup_supplier	0.000034
event_dayofweek	0.000021

Table 4. Feature importance rankings for production-impact prediction

Table 4 presents the leading predictors identified through the aggregated tree based feature importance analysis. The results indicate that missed_bytes was overwhelmingly the dominant predictor, accounting for approximately 99.8% of the aggregated impurity-based importance of the tree model. The remaining variables, including ossec_alert_level, successful_logins, paket_rate, duration_s, byte_rate, and patch_level, contributed substantially smaller proportions of the total importance.

The prominence of missed_bytes suggests that network-traffic irregularities or disruptions in data transmission are strongly associated with the magnitude of production impact in the analysed dataset. Other variables, such as security-alert levels, login activity, network packet rates, connection duration, and patch-level compliance, may provide additional information about the operational and cybersecurity conditions associated with production disruption.

Nevertheless, the exceptionally high importance assigned to missed_bytes requires careful interpretation. Feature importance identifies predictive contribution within the fitted model; it does not establish a causal relationship. More importantly, if missed-byte measurements are recorded during or after a production disruption, the variable may represent a contemporaneous indicator of an event that has already occurred rather than an antecedent warning signal.

Therefore, the variable's measurement timing should be verified before describing the model as an early-warning system. The sensitivity analysis in Table 5 provides an additional assessment of the extent to which model performance depends on this predictor.

Figure 1 illustrates the comparative predictive performance of the eight regression algorithms using their respective test-set coefficients of determination. The visualization provides an overall assessment of how effectively the models explain variations in production impact among suppliers excluded from the training dataset.

The Gradient Boosting Regressor achieved the highest R^2 value of 0.7679, followed closely by Random Forest at 0.7675 and XGBoost at 0.7660. These results demonstrate that ensemble-based tree algorithms provided the strongest predictive performance among the evaluated approaches. The Artificial Neural Network achieved

an R^2 of 0.7626, while Ridge, Linear, and Lasso Regression produced similar values of approximately 0.7526. Support Vector Regression recorded the lowest R^2 of 0.7391.

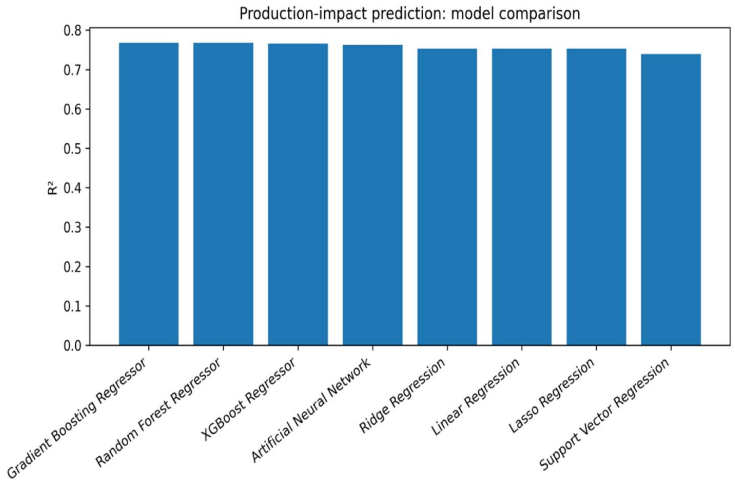

Figure 1. Test-set R^2 comparison across the eight regression algorithms

The relatively narrow performance differences among the leading models indicate that several algorithms were capable of learning useful predictive relationships from the available supplier and cybersecurity indicators. However, Gradient Boosting provided the best observed balance of prediction accuracy and explained variance. The figure therefore supports its selection as the primary model for subsequent diagnostic and sensitivity analyses. Since repeated grouped validation was not performed, the ranking should be understood as a comparison on this particular held-out test set rather than a definitive ranking across all possible supplier populations.

6.2 Feature importance

The leading predictor was missed_bytes, with approximately 99.8% of the tree model's aggregated impurity-based importance. This exceptionally dominant contribution must be interpreted cautiously. If missed_bytes is measured during or after the disruption, it is a contemporaneous indicator rather than a valid early-warning predictor.

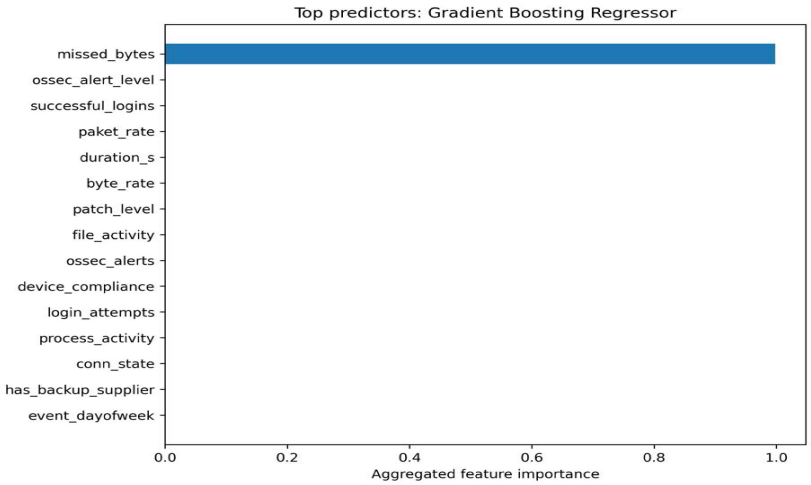

Figure 2. Leading predictors according to aggregated tree-based feature importance

Figure 2 presents the relative importance of the leading predictors used by the tree based model for estimating production disruption. The visualization demonstrates a highly concentrated importance distribution, with missed_bytes accounting for approximately 99.8% of the aggregated impurity based importance.

The remaining predictors, including OSSEC alert levels, successful logins, packet rate, session duration, byte rate, patch level, file activity, device compliance, login attempts, and process activity, contribute comparatively small proportions. This distribution suggests that the fitted model relies predominantly on one network related variable when generating its predictions.

The dominance of missed_bytes has two important implications. First, it highlights the potential importance of network-level indicators in estimating production disruption. Second, it raises a methodological concern regarding the timing and availability of this variable. If missed byte information is generated as a consequence of an incident or its operational impact, its high importance may reflect the model's ability to recognize an already developing disruption rather than predict one in advance.

Consequently, Figure 2 should be interpreted as a description of the model's internal predictive structure rather than evidence that missed bytes independently cause production losses. The results also motivate the sensitivity analysis presented in Figure 5 and Table 5.

6.3 Prediction diagnostics

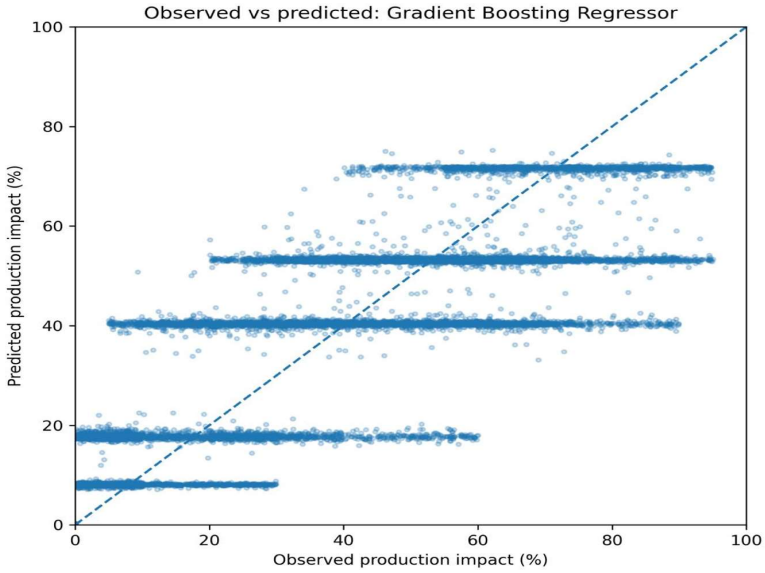

Figure 3. Observed versus predicted production impact on held out suppliers

Figure 3 evaluates the predictive behaviour of the selected Gradient Boosting Regressor by comparing observed production impact percentages with the corresponding model predictions for the held out suppliers. The horizontal axis represents the observed production impact, while the vertical axis represents the predicted values. The diagonal reference line represents perfect prediction, where predicted values are identical to observed values.

Observations concentrated near the diagonal line indicate relatively accurate predictions, whereas points located farther from the reference line indicate larger prediction errors. The overall alignment of the observations with the reference relationship provides a visual assessment of the model's capacity to reproduce the variation in production impact across previously unseen suppliers.

Deviations from the diagonal line may indicate prediction errors associated with differences in supplier characteristics, cybersecurity conditions, network activity, and incident severity. In particular, the presence of zero impact observations and the wide range of positive production-impact values may introduce heterogeneous prediction patterns. The model may also exhibit different levels of accuracy across low , moderate and high impact incidents.

The figure complements the quantitative evaluation reported in Table 1 by showing the relationship between actual and estimated values rather than relying exclusively on aggregate error metrics. It is particularly useful for identifying potential systematic overprediction or underprediction. However, the precise extent of such patterns should be established from the plotted observations and not inferred solely from the overall R^2 value

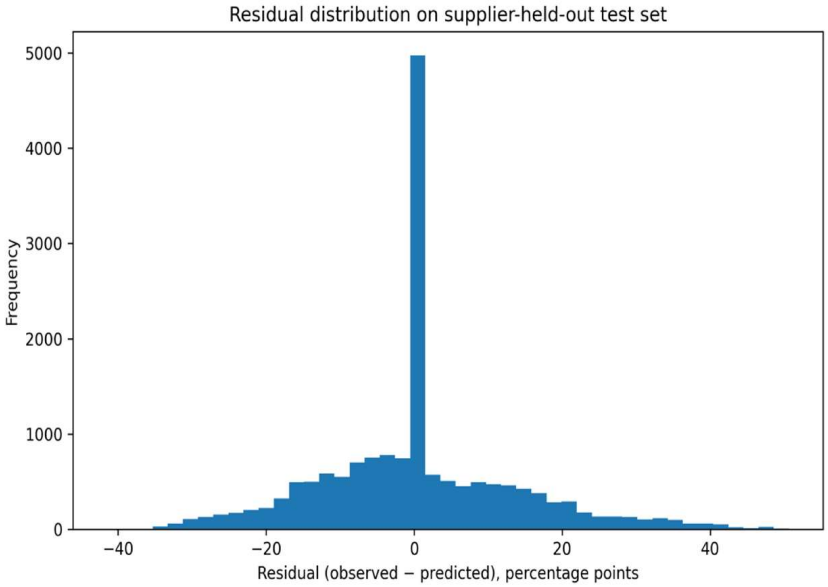

Figure 4. Residual distribution for the selected model

Figure 4 presents the distribution of residual errors generated by the selected Gradient Boosting Regressor. Residuals are calculated as the difference between observed and predicted production impact percentages. The residual distribution provides information about the magnitude, direction, and variability of prediction errors.

Residuals concentrated around zero indicate that the model does not consistently overestimate or under estimate production impact across the test observations. A relatively narrow distribution would indicate that prediction errors are generally small, whereas a wider distribution would indicate greater variability in model performance. Skewness, heavy tails, or multiple concentrations may indicate that prediction errors differ across production-impact levels or supplier groups.

This diagnostic is particularly relevant because the target variable is bounded between 0% and 95% and includes a substantial proportion of zero-impact observations. Such characteristics may produce non-uniform residual behaviour, particularly when the model attempts to estimate very low or very high production impact values.

The residual analysis therefore complements MAE, MSE, and RMSE by examining the structure of prediction errors. It also provides a basis for determining whether further modelling approaches, such as zero inflated or two stage regression, may improve performance. The residual distribution should not, by itself, be treated as proof that all regression assumptions are satisfied.

6.3 Sensitivity analysis

Specification	MAE	MSE	RMSE	R^2
With missed_bytes	9.210	178.283	13.352	0.7679
Without missed_bytes	9.668	197.952	14.070	0.7423

Table 5. Sensitivity analysis of Gradient Boosting with and without missed_bytes

Table 5 evaluates the robustness of the selected Gradient Boosting Regressor by comparing two predictor configurations: the original specification containing missed_bytes and an alternative specification in which this variable was excluded. This comparison is important because the feature-importance analysis identified missed-byte measurements as the dominant predictor.

The original model achieved an MAE of 9.210, MSE of 178.283, RMSE of 13.352, and R^2 of 0.7679. After removing missed_bytes, MAE increased to 9.668, MSE increased to 197.952, RMSE increased to 14.070, and R^2 declined to 0.7423. Thus, excluding the variable resulted in an increase in prediction error and a reduction of approximately 2.56 percentage points in explained variance.

These findings demonstrate that the predictive performance of the selected model depends materially on the information provided by missed_bytes. Nevertheless, the alternative model retained an R^2 of 0.7423, suggesting that the remaining supplier, network, and cybersecurity indicators also contain substantial predictive information.

From an operational perspective, the comparison has important implications for deployment. If missed-byte measurements are available before production impact occurs, their inclusion may improve predictive accuracy. Conversely, if they become available only during or after the disruption, the model without this variable represents a more conservative specification for advance warning applications. The distinction between retrospective impact estimation and prospective disruption prediction should therefore be explicitly maintained.

Figure 5 compares the predictive performance of the Gradient Boosting Regressor under two predictor configurations, namely the original model containing missed_bytes and the alternative model excluding this variable. The comparison considers MAE, MSE, RMSE, and R^2. Removing missed_bytes increased MAE from 9.210 to 9.668 and RMSE from 13.352 to 14.070 percentage points. The R^2 declined from 0.7679 to 0.7423. These findings demonstrate that the model derives a substantial proportion of its predictive information from this network-related variable. Nevertheless, the alternative specification retains considerable explanatory performance, indicating that other supplier and cybersecurity indicators also contribute to production-impact estimation. The results emphasize the importance of verifying whether missed_bytes is available before production disruption occurs.

Figure 6 illustrates the relationship between predicted production impact percentages and residual errors for the 16,996 held-out observations. Residuals are calculated as observed minus predicted production impact; therefore, positive residuals indicate underprediction, whereas negative residuals indicate overprediction. The horizontal reference line represents zero prediction error.

The residuals are distributed around the zero-reference line, with a mean residual of approximately 0.43 percentage points and a residual standard deviation of approximately 13.35 percentage points. The spread of residuals varies across predicted-impact levels, with relatively wider dispersion in several intermediate and higher prediction ranges. This suggests that prediction accuracy is not uniform across the production impact spectrum. The visible vertical concentrations also reflect the discrete prediction values generated by the tree based ensemble model. Overall, the figure indicates that the model captures the general variation in production impact but retains substantial event level prediction errors, particularly for some observations with larger production consequences.

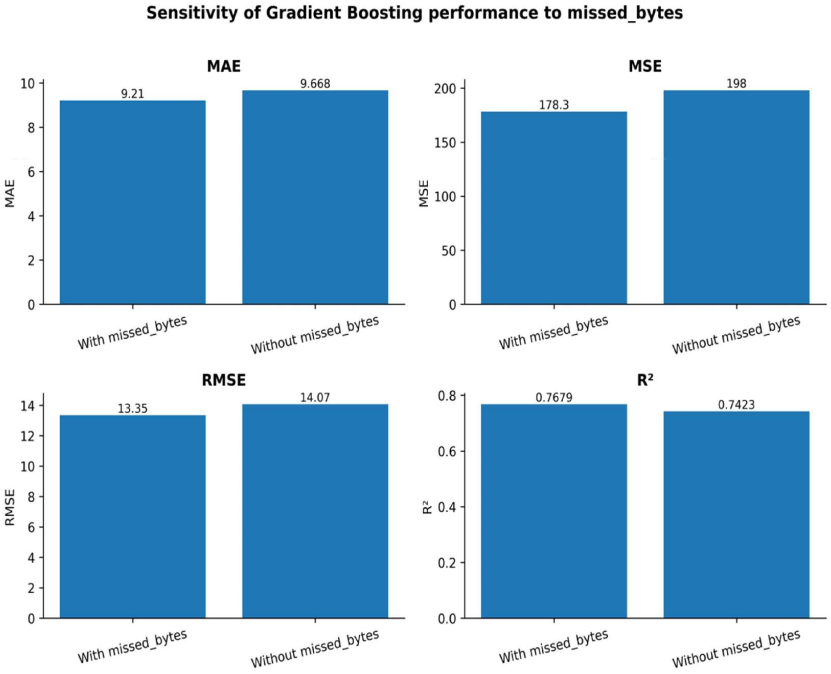

Figure 5 – Sensitivity comparison

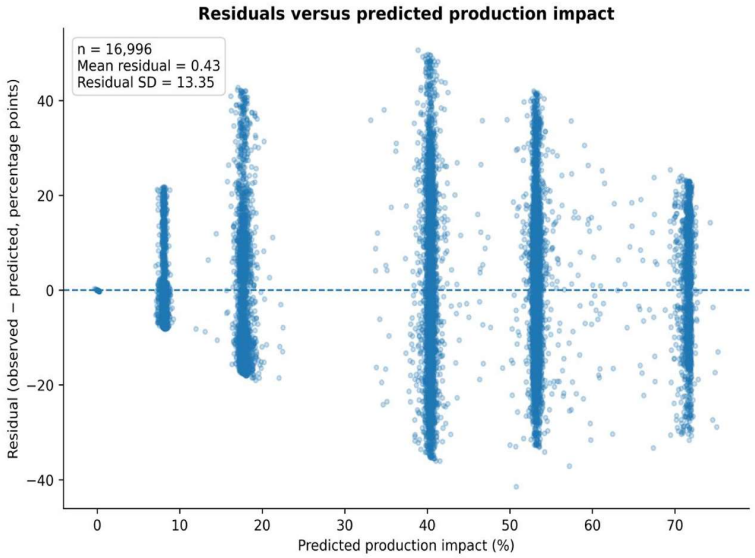

Figure 6. Residuals versus predicted production impact

Figure 7 presents the SHAP summary plot, which explains the contribution of individual predictors to the Gradient Boosting model's production-impact estimates. Each point represents an observation, while its horizontal position indicates the SHAP value. Positive SHAP values increase the model's predicted production impact, whereas negative values reduce it. Point colour represents the corresponding feature value, with blue indicating lower values and pink indicating higher values.

The figure identifies missed_bytes as the most influential predictor by a substantial margin. Higher values of

this variable are generally associated with positive SHAP contributions, indicating an increase in estimated production impact, whereas lower values tend to contribute negatively. Other variables, including ossec_alert_level, successful_logins, paket_rate, file_activity, and byte_rate, exhibit considerably smaller contributions. This pattern is consistent with the previously reported tree-based feature importance analysis. The results indicate that the fitted model relies predominantly on network traffic information, while the remaining security and operational variables provide relatively limited additional predictive contributions. These SHAP results explain model behaviour and should not be interpreted as evidence of causal effects.

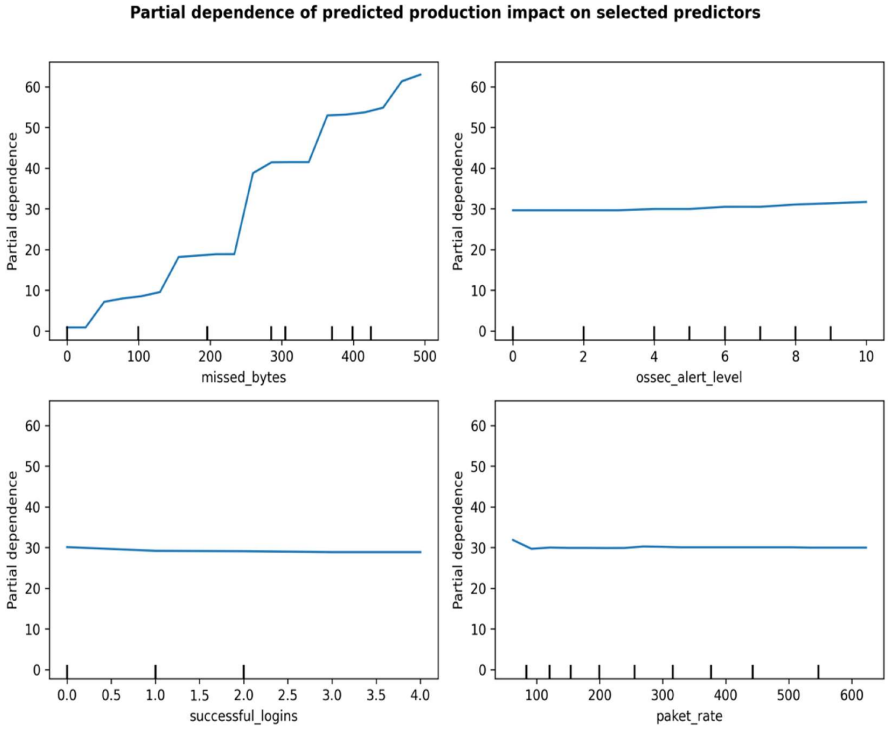

Figure 8. Partial dependence plots of selected production impact predictors

Figure 8 presents partial dependence plots for four leading predictors: missed_bytes, ossec_alert_level, successful_logins, and paket_rate. Each curve illustrates the average predicted production impact as the selected feature varies, while the effects of other predictors are averaged over the evaluation observations.

The partial dependence curve for missed_bytes demonstrates a pronounced nonlinear and stepwise increasing relationship with predicted production impact. In particular, predicted impact rises substantially across certain ranges of missed-byte values, indicating that the model associates higher network disruption measurements with greater production consequences. In contrast, the curves for OSSEC alert level, successful logins, and packet rate are comparatively flat. This suggests that these variables have much smaller marginal effects on the model's average predictions than missed_bytes.

These findings reinforce the SHAP analysis by demonstrating that the model's predicted production impact is particularly sensitive to changes in missed-byte measurements. However, partial dependence represents an average model response rather than a causal relationship. Correlations among predictors and the timing of measurement should therefore be considered when interpreting the curves.

Figure 9 compares the mean observed and predicted production impact percentages across five categories defined using the observed target values: zero impact, low impact, moderate impact, high impact, and very high impact. The number of observations in each category is displayed above the corresponding bars

The results reveal different prediction patterns across impact categories. For zero impact observations, the mean predicted impact was approximately 0.15%, indicating that the model estimated values close to zero. For the low and moderate impact categories, the model overestimated the observed mean impact: the observed and predicted means were 10.21% and 18.11% for the low-impact category and 37.87% and 42.76% for the moderate impact category, respectively.

Conversely, the model underestimated production impact in the high and very high impact categories. For high-impact observations, the observed mean was 61.38%, compared with a predicted mean of 51.97%. The difference was more pronounced for very high impact observations, where the observed mean was 84.03%, whereas the predicted mean was 61.10%.

These findings suggest that the model tends to moderate extreme predictions, overestimating some low impact events while underestimating the most severe disruptions. This behaviour is important for supply chain risk management because underestimation of very high impact incidents could lead to insufficient preparation for severe operational consequences. Future modelling should therefore investigate category specific prediction errors, weighted regression, and two stage modelling strategies.

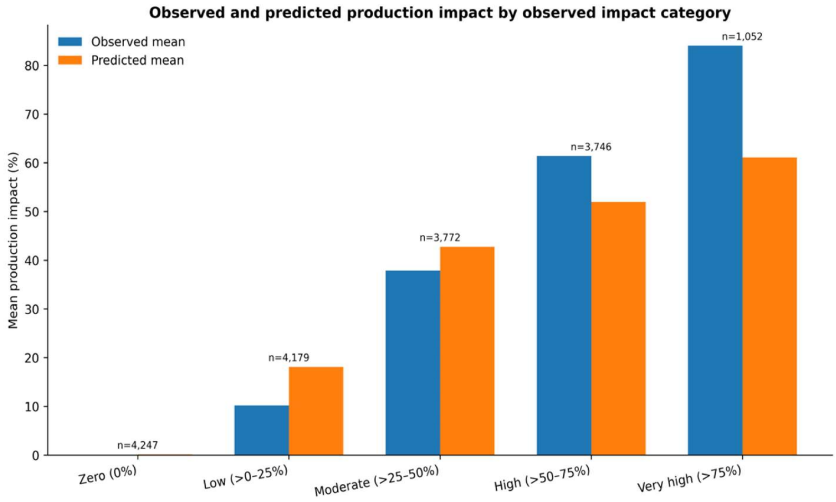

Figure 9. Observed and predicted production impact across impact categories

8. Discussion

Taken together, Tables and Figures demonstrate the potential of machine-learning regression techniques for estimating production disruption associated with cybersecurity incidents in supply chains. The comparative evaluation identifies Gradient Boosting as the strongest performing algorithm, achieving an R^2 of 0.7679 and an RMSE of 13.352 percentage points under supplier grouped validation. The close performance of Random Forest and XGBoost further indicates the suitability of ensemble learning for capturing complex relationships between supplier characteristics, network activity, cybersecurity indicators, and operational outcomes.

The validation protocol strengthens the reliability of the performance assessment by ensuring that suppliers in the test dataset were completely excluded from model training. This reduces the possibility that the reported performance is attributable simply to the model recognizing supplier specific characteristics already encountered during training. Nevertheless, because the validation was not chronological, the findings establish cross-supplier generalization rather than future-event forecasting.

The feature importance and sensitivity analyses identify missed_bytes as a particularly influential predictor. Its removal reduced the model's R^2 from 0.7679 to 0.7423, demonstrating that the model retains substantial

predictive capability even without this variable. This result is important for operational deployment because it distinguishes a model that estimates disruption using contemporaneous network information from one that may be suitable for advance warning.

Overall, the findings support the use of machine learning for supply chain production impact assessment while emphasizing the need for temporally valid predictors, supplier independent validation, and additional robustness testing. Future research should investigate chronological validation, repeated supplier-grouped cross-validation, hyperparameter optimization, and two stage modelling to address the substantial concentration of zero-impact observations.

9. Limitations

Several limitations should be considered when interpreting the findings. First, the supplier grouped train test split establishes generalisation to previously unseen suppliers but does not evaluate temporal generalisation. Because the partition was not chronological, the reported performance does not demonstrate the ability to forecast future disruption events. Second, the models capture predictive associations rather than causal relationships; the estimated coefficients and feature contributions should not be interpreted as causal effects of individual predictors on production impact.

Third, hyperparameter optimisation was limited and repeated supplier grouped cross validation was not performed, so the ranking of models and the absolute performance figures may be sensitive to alternative tuning and resampling schemes. Fourth, Support Vector Regression was trained on a random 5,000 row subset for computational reasons, whereas all other algorithms used the full training set, which may have placed SVR at a relative disadvantage.

Fifth, the target variable is bounded between 0 and 95 % and exhibits a substantial concentration of zero-impact observations (approximately 25 %). A single stage regression model does not fully exploit this structure; a two-stage approach that first classifies zero versus positive impact and then regresses on the positive cases could improve calibration and reduce error in low impact regimes. Finally, the dominant predictor missed_bytes may incorporate information recorded during or after an incident. Without explicit verification of its measurement timing relative to the onset of production loss, claims that the model functions as a true prospective early warning system cannot yet be substantiated.

10. Conclusion

This research formulated the estimation of cybersecurity related production disruption as a supervised regression problem and systematically compared eight machine learning algorithms under a supplier grouped validation regime. Gradient Boosting Regressor delivered the strongest observed performance, explaining approximately 77 % of the variance in production impact percentage on held out suppliers (MAE $\approx$ 9.2 percentage points, RMSE $\approx$ 13.4 percentage points). Ensemble tree based methods consistently outperformed linear and kernel based alternatives, indicating that non-linear interactions among supplier attributes, network-traffic statistics, and cybersecurity indicators are important for accurate estimation.

Explainability analyses confirmed that a small set of network and security features, led by missed_bytes, accounts for most of the predictive signal, while sensitivity experiments showed that the model retains useful accuracy even when this dominant feature is removed. These results demonstrate that multi-source operational and cybersecurity data can support supplier independent assessment of production impact.

At the same time, the non chronological validation design and unresolved questions about predictor timing limit the immediate translation of the framework into an operational early warning system. Future work should therefore prioritise chronological validation, repeated grouped cross validation with uncertainty quantification, exhaustive hyperparameter search, and two stage modelling strategies that explicitly handle zero inflated outcomes. With these extensions, the approach developed here can evolve from a retrospective impact estimator into a practical component of proactive supply chain risk management.

Reference

[1] Widayanti, R., Setiyowati, H., Yusup, M., Rodriguez, M. (2026). Predicting supply chain risks using machine learning for resilient operations. *ADI Journal on Recent Innovation (AJRI), 7*(2), 137–148.

[2] system: Evidence from US semiconductor industry. *Journal of Advanced Computing Intelligent Systems, 3*(11), 21–35.

[3] Sarker, B. (2025). Leveraging machine learning for supply chain disruption prediction and risk assessment in US industries. *International Journal of Applied Mathematics, 38*(10s), 2382–2410.

[4] Fareed, S. M. (2026). *Machine learning based supplier risk prediction and resilience scoring for strategic sourcing in US manufacturing supply chains* (SSRN Working Paper No. 6495779). SSRN.

[5] Ivanov, D., Sokolov, B., Solovyeva, I., Dolgui, A., Jie, F. (2016). Dynamic recovery policies for time-critical supply chains under conditions of ripple effect. *International Journal of Production Research, 54*(23), 7245–7258.

[6] Ivanov, D. (2017). Simulation-based ripple effect modeling in the supply chain. *International Journal of Production Research, 55*(7), 2083–2101.

[7] Rozhkov, M., Rozhkov, M. (2018). Capacity production-inventory planning. *IFAC-PapersOnLine, 51*(11), 1448–1452.

[8] Dolgui, A., Ivanov, D., Sokolov, B. (2018). Ripple effect in the supply chain: An analysis and recent literature. *International Journal of Production Research, 56*(1–2), 414–430.

[9] Ivanov, D., Dolgui, A., Sokolov, B. (2018). Scheduling of recovery actions in the supply chain with resilience analysis considerations. *International Journal of Production Research, 56*(19), 6473–6490.

[10] Ivanov, D., Dolgui, A. (2019). Low certainty need (LCN) supply chains: A new perspective in managing disruption risks and resilience. *International Journal of Production Research, 57*(15–16), 5119–5136.

[11] Ivanov, D. (2019). Disruption tails and revival policies: A simulation analysis of supply chain design and production-ordering systems in the recovery and post disruption periods. *Computers Industrial Engineering, 127*, 558–570.

[12] Hosseini, S., Ivanov, D., Dolgui, A. (2020). Ripple effect modeling of supplier disruption: Integrated Markov chain and dynamic Bayesian network approach. *International Journal of Production Research, 58*(11), 3284–3303.

[13] Llaguno, A., Mula, J., Campuzano Bolarin, F. (2021). State of the art, conceptual framework and simulation analysis of the ripple effect on supply chains. *International Journal of Production Research, 60*(6), 2044–2066.

[14] Dolgui, A., Ivanov, D., Dolgui, A. (2021). Ripple effect and supply chain disruption management: New trends and research directions. *International Journal of Production Research, 59*(1), 102–109.

[15] Xu, S., Zhang, X., Feng, L., Yang, W. (2020). Disruption risks in supply chain management: A literature review based on bibliometric analysis. *International Journal of Production Research, 58*(11), 3508–3526.

[16] Ali, S. M., Bari, A. B. M. M. M., Rifat, A. A. M. M., Alharbi, M., Choudhary, S., Luthra, S. (2022). Modeling supply chain disruption analytics under insufficient data: A decision support system based on Bayesian hierarchical approach. *International Journal of Information Management Data Insights, 2*(2), 100121.