Sciences and Engineering Research
AIJSER VOL 9 NO 1 (2026) E-ISSN 2641-0311 P-ISSN 2641-0303
Available online at www.acseusa.org
Journal homepage: https://www.acseusa.org/journal/index.php/aijser
Published by American Center of Science and Education, USA

Engineered risk features for trade-exposure classification: a gradient boosting framework validated on synthetic firm data

Netranand Pathak [(a)] *Rhythm Rakesh* [(b)1] *Shashikant Rai* [(c)] *Chelpuri Sumithra Devi* [d)]

[(a)] *Research Scholar, Department of Management Studies, Indian Institute of Information Technology Allahabad, Uttar Pradesh, India; E-mail: rsm2023005@iiita.ac.in*
[(b)] *Research Scholar, Department of Management Studies, Indian Institute of Information Technology Allahabad, Uttar Pradesh, India; E-mail: rsm2025002@iiita.ac.in*
[(c)] *Associate Professor, Department of Management Studies, Indian Institute of Information Technology Allahabad, Uttar Pradesh, India; E-mail: shashikant@iiita.ac.in*
[(d)] *Student, Department Of Information Technology, Indian Institute of Information Technology Allahabad, Uttar Pradesh, India; E-mail: iec2023112@iiit.ac.in*

ARTICLE INFO

Article History:

Received: 21st June 2026
Reviewed & Revised: 21st June 2026 to 8th November 2026
Accepted: 10th November 2026
Published: 13th November 2026

Keywords:

Trade Exposure, Supply Chain Risk, Firm Heterogeneity, Feature Engineering, Gradient Boosting, Risk Classification, Indian Exporters, Synthetic Benchmark

JEL Classification Codes:

C45, C53, F14, F15, M11

ABSTRACT

Realignments in global trade relationships affect exporting firms unevenly, yet risk assessment commonly relies on sector aggregates that conceal firm-level differences in exposure and adaptive capability. Operational measures that combine trade exposure with supply-chain capability for individual firms remain scarce. This study examines whether a domain-engineered representation of firm-level trade vulnerability improves risk classification and whether the improvement depends on learner capacity. Five composite indicators (Trade Concentration, Export Vulnerability, Diversion Exposure, Resilience Gap and Import Dependency) are constructed from six trade and capability variables grounded in the Indian trade literature. The study employs a synthetic dataset of 10,000 firms with a specified risk structure, ten random seeds and an ablation of raw, engineered and combined feature sets across Decision Tree, Logistic Regression, Random Forest and Gradient Boosting classifiers. The results show that Gradient Boosting classifies firms into Low, Medium and High risk with 84.44 per cent accuracy, a macro-F1 of 0.8392 and one-versus-rest AUC values between 0.9168 and 0.9828; 99.74 per cent of misclassifications fall between adjacent categories. The macro-F1 gain from the engineered indicators declines from 0.0528 for the Decision Tree to 0.0370 for Logistic Regression, 0.0186 for Random Forest and 0.0062 for Gradient Boosting (Holm-adjusted $p = 0.106$). The five indicators alone perform within 0.0030 macro-F1 of the full eleven-variable set. The findings of this study suggest that domain feature engineering and model capacity act as partial substitutes and that export vulnerability, diversion exposure and import dependency are the dominant dimensions of firm-level trade risk within the benchmark.

© 2026 by the authors. Licensee ACSE, USA. This article is an open access articledistributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

INTRODUCTION

When two major trading partners strengthen their bilateral ties, the impact on exporters in third countries is rarely even. Some lose orders rapidly; others absorb the pressure by diversifying their destinations or by making their sourcing more flexible; a third group finds out only later that their upstream suppliers have been reconfigured around new trade patterns. This is a significant commercial issue for Indian firms exporting to the European Union and the United States, and one that traditional trade analysis addresses poorly, as inter-industry differences are often smaller than intra-industry differences (Bernard et al., 2012). Risk exposure can be difficult to see at the sector level or in aggregate export statistics. The key is the distribution: which firms in a sector are concentrated in the affected destinations, which are not flexible enough to reroute, and which rely on imported inputs whose supply lines are also being rerouted? The empirical evidence on firm-level heterogeneity in trade participation is well documented (Bernard et al., 2007, 2012; Melitz, 2003), and supply-chain theory has long maintained that vulnerability depends not just on exposure, but also on the ability to withstand and respond to disruption (Pettit et al., 2010; Wagner & Bode, 2006). There is little overlap between these two literatures at the level of measurement. While trade economics provides exposure constructs and supply-chain management provides resilience and flexibility constructs, there are still few operational representations that integrate the two into constructs that can be calculated for individual firms, and measurement is still recognised as the constraint in work on global supply-chain risk (Baldwin & Freeman, 2022; Bhattacharjee et al., 2026; Dev et al., 2022). This study aims to fill this void by designing a small set of engineered risk indicators that explicitly incorporate exposure and adaptive capability, and by examining the impact of using such a representation of firm risk on classification performance. Six trade and capability variables are used

[1]*Corresponding author: ORCID ID: 0009-0002-0476-2304*
© 2026 by the authors. Hosting by ACSE. Peer review under responsibility of American Center of Science and Education, USA.
https://doi.org/10.46545/fze96p33

To cite this article: Pathak, N., Rakesh, R., Rai, S., & Devi, C. S. (2026). Engineered risk features for trade-exposure classification: a gradient boosting framework validated on synthetic firm data. *American International Journal of Sciences and Engineering Research*, *9*(1), 12-23. https://doi.org/10.46545/fze96p33

to build five composite indicators, and a Gradient Boosting classifier is employed to classify firms into Low, Medium and High risk categories. Three other learners with varying complexity are provided as comparators, enabling a second question to be answered: how far does the benefit of domain-informed feature construction depend on the sophistication of the model consuming it? Testing the proposed representation requires a setting in which the risk structure can be defined beforehand. The study therefore examines the indicators on a controlled synthetic firm data set, enabling the role of the representation to be systematically explored across learners. The constructs are based on the empirical trade and supply-chain literature, and the benchmark tests the value of the engineered representation. The paper makes four contributions. First, it maps broad ideas of trade exposure into a structured and interpretable representation of the vulnerability of the firm, instead of a single exposure ratio. Second, it makes explicit the interaction between exposure and adaptive capability: two firms with the same market exposure have different risk positions if they have different flexibility and input dependence, and the indicators reflect this interaction. Third, it investigates whether the representation contains discriminatory information and demonstrates that the benefit of explicitly constructing features is a systematic function of model capacity, which is relevant to the allocation of effort between domain features and model complexity. Fourth, it shows a controlled evaluation path for risk representations when outcome labels are not observable. The analysis is organised around four questions.

RQ1: Does a domain-engineered representation improve firm-level risk classification relative to raw trade variables?
RQ2: Does the size of that improvement depend on the capacity of the learner?
RQ3: Can the engineered indicators alone provide a compact representation, or are the underlying raw variables also required?
RQ4: Are these patterns stable across alternative specifications of the risk structure?

LITERATURE REVIEW

Trade policy and trade diversion

Since Viner (1950/2014), trade diversion, the redirection of trade flows as a result of preferential arrangements between other parties, has been a key topic in trade economics and remains the main lens for evaluating preferential market access (Fugazza & Nicita, 2013). Its effects fall not only on trade volumes but also on the competitive position of firms, which shifts according to the distribution of their market presence. The effects of trade-policy uncertainty on investment and trade flows are well documented (Caldara et al., 2020; Handley & Limão, 2017), and the impact of policy shocks on real firm outcomes has been found to pass through exposure channels (Pierce & Schott, 2016). This literature shows that exposure matters and that it is unevenly distributed; hence, there is a need for firm-level representations that can identify vulnerable firms before a realignment occurs (Baldwin & Freeman, 2022).

Firm heterogeneity in trade exposure

The heterogeneous-firms literature has established that exporters are not randomly drawn from the population of firms but differ from non-exporters and from one another in productivity, size and market participation, and that trade shocks therefore reallocate activity within industries, but not uniformly (Melitz, 2003). Empirical studies have found that only a small share of firms export, that exporting firms differ widely in the destinations they serve and the products they export, and that variation among exporting firms is generally larger than variation among non-exporting firms (Bernard et al., 2007, 2012). Indian evidence is similar: export participation is highly uneven, even within narrowly defined industries; firms' positions in global value chains differ widely; and export intensity is driven by firm characteristics like entrepreneurial orientation and business-group affiliation, rather than by industry membership (Banerjee et al., 2022; Ghosh & Sinha Roy, 2018; Reddy & Sasidharan, 2021). The implication for risk assessment is direct and under-exploited: if exposure is more variable within a sector than between sectors, then sector-level risk classification can misallocate attention: low-exposure firms in nominally risky sectors can be flagged, while concentrated, inflexible firms in nominally safe sectors can go unflagged. The translation of this insight into a usable screening representation, one defined on firm characteristics instead of industry membership, remains largely under-researched.

India as the empirical context

The present indicators were developed in the context of India. Indian firms are not equally involved in international trade, and the exposure of firms is determined not only by the destination of their exports but also by the intermediate inputs that they import. Domestic product growth and firm performance increased following tariff liberalisation (Goldberg et al., 2010; Topalova & Khandelwal, 2011); for Indian manufacturers, the productivity impact of input tariffs was found to be linked to institutional conditions across states, with the gains being felt through input costs and markups, not just prices (Ahsan, 2013; Arneja & Sharma, 2026; De Loecker et al., 2016). Recent studies have correlated participation in the global value chain with productivity and markups, and have found that participation is limited by financial capacity and reduced by economic-policy uncertainty (Reddy & Sasidharan, 2021, 2024; Reddy et al., 2025). Three aspects of this evidence are relevant to risk representation. First, the upstream channel is not a secondary channel: if the firm depends on inputs from abroad, a change in the sourcing pattern can impact the firm even if the order book remains unchanged. Second, the European Union and the United States are the relevant destination markets in the trade-diversion scenario analysed here, and changes in the trade relationship between these two markets can also impact firms through upstream sourcing networks, meaning exposure can occur on either side of the firm's trade balance. Third, preferential arrangements benefit only those firms that discover and claim them; they are not always used, because doing so is difficult and expensive (Fugazza & Nicita, 2013). Because of limited engagement with the agreement landscape, the benefits on the output side may be lower than they could be, and the visibility of sourcing adjustments may be limited. All of these observations suggest that destination concentration, export

intensity, import dependence, trade-policy engagement and operational flexibility should be treated as separate risk dimensions and not as a single aggregate exposure ratio. The dimensions themselves are not characteristics of any particular national market but of the level of trade participation by firms, although their empirical distributions vary across economies.

Supply-chain risk and vulnerability

The supply-chain management literature approaches the same problem from the opposite angle and considers vulnerability to be a characteristic of the network structure and firm capability, not of trade policy. Wagner and Bode (2006) empirically found that vulnerability is not a function of exposure in the abstract, but of specific structural characteristics, namely dependence on individual customers and suppliers and the degree of sourcing concentration. Craighead et al. (2007) demonstrated that the severity of disruption is a function of both network design attributes and the mitigation measures that a firm has in place, so that the same shock can have different effects on different firms. Sourcing structure affects this directly: single sourcing concentrates exposure in one place, whereas multiple sourcing spreads it across several, at a cost (Namdar et al., 2018). Further studies have examined intertwined supply networks, where propagation and survivability are linked to structural interdependence (Ivanov & Dolgui, 2020), and quantitative reviews have mapped the propagation and survivability mechanisms that have been modelled (Hosseini et al., 2019). Exposure alone is therefore not a sufficient indicator of risk. While there is clear evidence that these constructs exist, much of the evidence is based on survey instruments administered to limited samples (Brandon-Jones et al., 2014; Wagner & Bode, 2006), leaving open how these constructs can be operationalised at scale from transactional data.

Exposure, flexibility and resilience

The moderating effect of capability is explicitly stated in the resilience literature. Pettit et al. (2010) defined resilience as the relationship between vulnerabilities and capabilities, stating that firms with a high ratio of capabilities to vulnerabilities are inefficient and those with a high ratio of vulnerabilities to capabilities are at risk. Jüttner and Maklan (2011) identified flexibility and visibility as the two key factors that influenced firms' reactions to the global financial crisis, while Chowdhury and Quaddus (2017) identified three components of resilience as a dynamic capability: proactive capability, reactive capability and supply-chain design. Measurement work has helped to move beyond a description of resilience as a property of the firm and into the realm of assessment: Ambulkar et al. (2015) developed and validated a scale for firm resilience to disruptions, and Brandon-Jones et al. (2014) differentiated resilience from robustness and demonstrated that both are conditioned by supply-base complexity and visibility. Firm-level evidence from natural disasters shows that supply-chain networks both amplify and dampen shocks depending on firm capability and network position (Kashiwagi et al., 2021; Todo et al., 2015), and that resilience is a costly investment decision in multi-tier chains (Grossman et al., 2023). The conceptual heart of the present study is that a highly exposed firm with a high degree of operational flexibility is a materially different risk proposition from an equally exposed firm without flexibility, and exposure and adaptability need to be represented together.

Feature engineering for risk assessment with domain knowledge

An analytical model can learn from a problem only in the way that the problem is represented. In applied predictive tasks, the generation of domain-specific features (ratios, differences and interactions) is a critical juncture at which substantive knowledge can be fed into predictive models (Domingos, 2012; Kuhn & Johnson, 2019), and it remains relevant for learners with high accuracy on raw tabular inputs (Grinsztajn et al., 2022). Heaton (2016) provides one of the few systematic evaluations and demonstrates that some transformations are easily synthesised by learners, while others are not. This shifts the practical question from whether feature engineering helps to which models need feature engineering. In a risk-screening context the argument is further bolstered by interpretability: an indicator such as export intensity weighted by operational rigidity can be explained to a manager and audited by a regulator, whereas a representation learned internally by a model can be more difficult to communicate and audit.

Machine learning for firm-level risk classification

Gradient Boosting constructs trees sequentially, each correcting the errors of its predecessors (Friedman, 2001). The approach is now standard for structured tabular problems (Chen & Guestrin, 2016) and is suitable for risk classification, where errors are concentrated in ambiguous regions: firms with moderate exposure and moderate flexibility are intrinsically harder to classify than clear cases. Tree ensembles have been shown to be competitive in credit risk benchmarking on structured firm-level data (Baesens et al., 2003; Lessmann et al., 2015), and they remain so against more recent alternatives (Grinsztajn et al., 2022; Shwartz-Ziv & Armon, 2022). The use of learners of varying complexity as comparators is not only for the purpose of model selection, but also because a simple additive model cannot represent an interaction unless it is explicitly provided, whereas a boosted ensemble can approximate an interaction internally. Comparing across learners thus shows how much of the value of a domain-informed representation is genuinely additional, rather than something that a sufficiently flexible model would have recovered unaided.

Interpreting risk features

For risk screening, the classification itself matters less than knowing which indicators drive it. Impurity-based importance measures in tree ensembles are known to be sensitive to variable characteristics (Strobl et al., 2007), and permutation-based measures can be difficult to interpret when features are correlated or derived from one another (Hooker et al., 2021; Molnar et al., 2022), a direct concern here, since composite indicators are by construction functions of shared underlying variables. Rudin (2019) argues that interpretable structure is preferable to post-hoc explanation where decisions carry consequences.

This supports building interpretability into the representation and also suggests that the attribution results in this context should be interpreted as evidence of ordering, not of the exact size of the contribution of each indicator.

Research gap

Two literatures meet here without having been brought into contact. Trade economics and supply-chain management provide well-developed concepts such as exposure, destination dependence, supplier dependence, upstream input dependence, flexibility and adaptive capability, but do not often capture these concepts as quantities that can be computed across a population of firms for screening purposes. Predictive modelling provides the tools to encode and test such constructs, offering an opportunity to explore whether the engineered representation offers predictive power over the raw variables and whether the predictive power is contingent upon the learner's capacity. Addressing this gap requires a setting in which the risk structure is specified rather than inferred, an explicit ablation of feature conditions, and comparison across learners of differing representational power. Figure 1 summarises the study framework, linking the domain-grounded constructs to the engineered representation, the controlled benchmark and the empirical tests.

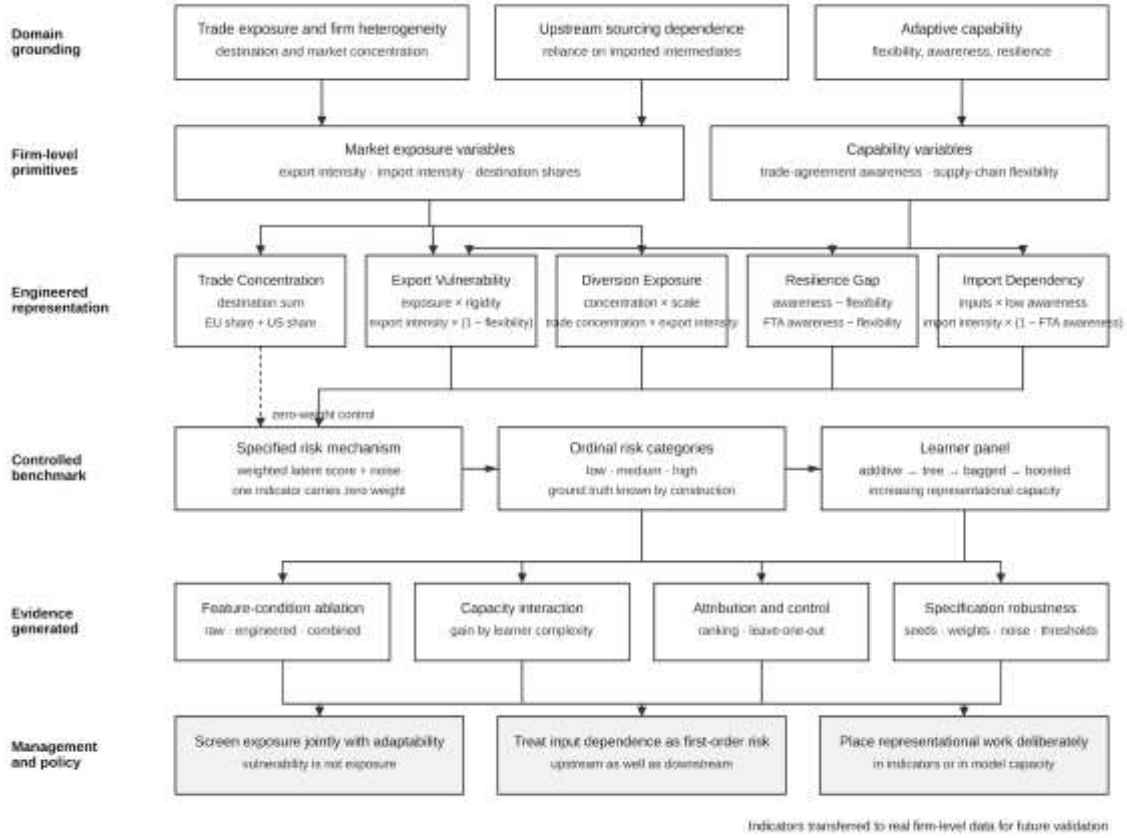

Figure 1. Research framework linking domain-grounded trade and supply-chain constructs, engineered risk representation, controlled benchmarking, and empirical evidence.

MATERIALS AND METHODS

Research design

The study does not estimate a trade relationship; it evaluates a risk representation. The question of interest is whether domain-informed indicators, calculated at the firm level, can help to accurately classify firms into risk categories, and whether this support varies by learner. The representation is tested on a synthetic firm data set where the risk structure is known a priori, as firm-level trade-risk labels are not directly observable. This allows the representation to be varied in a controlled way, while keeping all other factors constant.

Firm-level variables and their source

Six variables are used to describe each synthetic firm and cover the areas of market exposure and operational capability: Export Intensity, Import Intensity, EU Market Share, USA Market Share, FTA Awareness Score and Supply Chain Flexibility. These variables reflect three conceptual channels defined in the literature reviewed in Section 2: destination concentration as an exposure channel, similar to the customer-dependence mechanism defined in the vulnerability literature (Baldwin & Freeman, 2022; Wagner & Bode, 2006); imported-input reliance as a separate upstream channel in the Indian context (Ahsan, 2013; Arneja & Sharma, 2026; Goldberg et al., 2010); and operational flexibility as the primary moderating capability (Chowdhury & Quaddus, 2017; Pettit et al., 2010). The dimensions are justified in the context of Indian trade and supply chains, and the benchmark distributions are specified to facilitate controlled evaluation. Four variables are drawn from the unit interval [0,1). The two destination-share variables are drawn from a narrower support to keep the level of the

derived Diversion Exposure indicator within the benchmark. Draws are uncorrelated. A sector label (Textiles, Pharmaceuticals, Automotive, Electronics) is assigned independently and is used for descriptive purposes only; it does not enter any model and serves as an internal control to ensure that the classifier is not relying on category membership.

The engineered risk indicators

The five composite indicators convert the raw dimensions into meaningful measures of trade vulnerability. Each represents a different mechanism, and three of the five explicitly combine an exposure term with a capability term, the representational idea central to the study. Trade Concentration is a measure of dependence on the two destination markets involved in the trade-diversion scenario. Export Vulnerability identifies firms that are highly export-oriented but have limited ability to redirect. Diversion Exposure approximates the proportion of activity directly in the path of a realignment and is the product of destination dependence and export scale. Resilience Gap is the gap between awareness of a risk and the ability to respond; a positive gap indicates a firm that is aware of its risk but cannot move quickly. Import Dependency identifies firms that depend on imported inputs but have not taken part in the trade-agreement environment, a channel that the Indian evidence suggests is material.

For firm i, with EI export intensity, II import intensity, EU and US destination shares, FTA the trade-agreement awareness score and SCF supply-chain flexibility, the indicators are

$$TC_i = EU_i + US_i \tag{1}$$

$$EV_i = EI_i \times (1 - SCF_i) \tag{2}$$

$$DE_i = TC_i \times EI_i \tag{3}$$

$$RG_i = FTA_i - SCF_i \tag{4}$$

$$ID_i = II_i \times (1 - FTA_i) \tag{5}$$

Risk categories

A latent Network Risk Score is formed as a weighted combination of Export Vulnerability (0.35), Diversion Exposure (0.30), Import Dependency (0.20) and the absolute Resilience Gap (0.15). The weights are analyst-specified and treated as part of the benchmark specification. Robustness is assessed through five alternative weight configurations reported in the supplementary material. Gaussian noise ($\sigma = 0.035$) introduces stochastic variation into the risk score, and the resulting score is divided at the 45th and 78th percentiles into Low, Medium and High risk, producing a 45 / 33 / 22 per cent distribution.

$$R_i = 0.35\,EV_i + 0.30\,DE_i + 0.20\,ID_i + 0.15\,|RG_i| \tag{6}$$

$$\tilde{R}_i = R_i + \varepsilon_i, \qquad \varepsilon_i \sim N(0, \sigma^2), \qquad \sigma = 0.035 \tag{7}$$

$$y_i = \begin{cases} 0, & \tilde{R}_i \le q_{0.45} \\ 1, & q_{0.45} < \tilde{R}_i \le q_{0.78} \\ 2, & \tilde{R}_i > q_{0.78} \end{cases} \tag{8}$$

Trade Concentration carries zero weight in Equation (6). Because it also enters Diversion Exposure, it is treated as a partial internal control; its standalone importance is therefore interpreted together with the leave-one-out analysis reported.

Modelling approach

The main classifier is Gradient Boosting, as it builds trees in stages, each correcting the errors of its predecessors, and is appropriate for a scenario where the hard cases are clustered in the middle of the risk distribution. It is trained for 180 boosting rounds, learning rate 0.08, maximum depth 5, minimum leaf size 20 and 85 per cent subsampling; all of these parameters are fixed rather than optimised by hyperparameter search. All models were estimated using scikit-learn (Pedregosa et al., 2011), and a stratified 75/25 split was used to maintain class proportions. Three comparators of differing capacity (Logistic Regression, a depth-limited Decision Tree and a Random Forest) were tested under the same conditions. Their role is not competitive: an additive model cannot represent an interaction unless it is supplied, whereas a boosted ensemble can approximate one internally, so the comparison shows how much of the value of the representation is genuinely additional. In all such experiments, ten fixed seeds (99–108) are used, each of which regenerates the entire synthetic population and a new stratified 75/25 split. The stability of the observed representation and learner-capacity patterns is evaluated using seed-level results; primary-model results and ten-seed estimates are reported separately. Full specifications, ten-seed tables and dependence diagnostics are provided in Supplementary Material A.

RESULTS

Overall classification performance and error structure

The framework classifies firms into risk groups with a high degree of accuracy. The primary Gradient Boosting model achieves 84.44 per cent accuracy against a majority-class baseline of 45.0 per cent; five-fold cross-validation gives 84.54 per cent ± 0.68 per cent, and the average accuracy across ten seeds is 84.62 per cent (SD = 0.71 percentage points). The small cross-validation band shows the stability of the performance across the folds of the primary configuration. Macro-F1 is 0.8392, balanced accuracy 0.8378 and Cohen's κ 0.7566. Precision, recall and F1 are presented in Table 1 for each risk category. The two categories with the greatest screening consequences, Low risk and High risk, are those separated most reliably by the model (recall of 0.900 and 0.846, respectively, compared with 0.768 for Medium risk). Performance is strongest at the extremes and weakest in the middle: the Medium category is the most difficult to distinguish because its firms combine moderate exposure with some capacity to absorb shocks.

Table 1. Per-class classification performance, Gradient Boosting model.

Metric	Low Risk	Medium Risk	High Risk
Precision	0.895	0.763	0.864
Recall	0.900	0.768	0.846
F1-Score	0.897	0.766	0.855
Support	1,125	825	550

For a screening application, the more consequential question is not how many errors occur but how damaging they are, and here the structure is favourable. The confusion matrix in Table 2 shows that of 389 misclassifications, 388 fall between adjacent categories and exactly one crosses the Low–High boundary, so 99.74 per cent of errors are adjacent-category errors. Figure 2 presents the confusion matrix. For screening purposes, errors are not equally consequential: a High-risk firm classified as Low risk is treated as the most consequential error because it can result in an exposed firm being overlooked, whereas adjacent-category errors represent lower-severity classification deviations.

Table 2. Confusion matrix for Gradient Boosting risk-level classification, test set (n = 2,500).

Actual \ Predicted	Low Risk	Medium Risk	High Risk
Low Risk	1,012	112	1
Medium Risk	119	634	72
High Risk	0	85	465

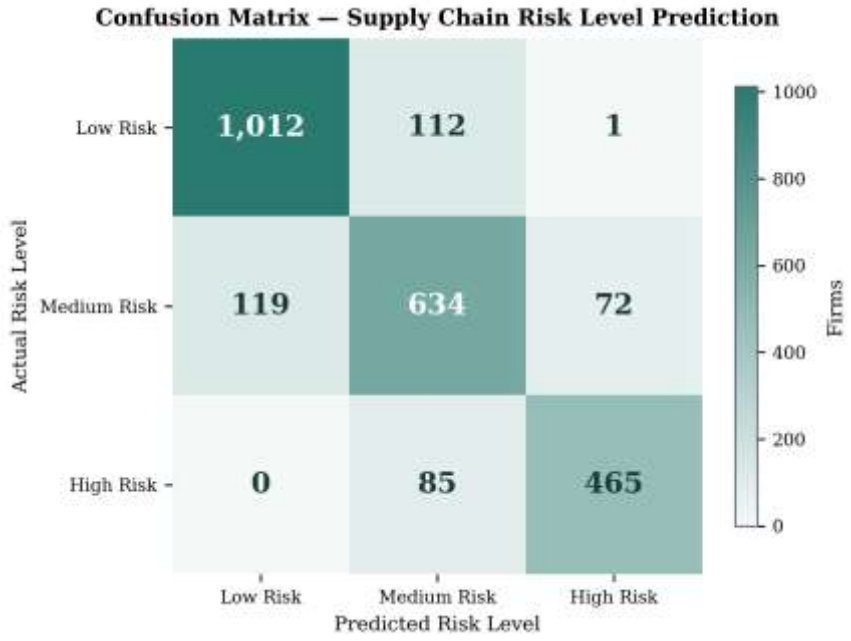

Figure 2. Confusion matrix for firm risk-level classification using Gradient Boosting (test set, n = 2,500).

Discrimination is consistent with this pattern. One-versus-rest AUC values are 0.9715 for Low risk, 0.9168 for Medium and 0.9828 for High. Discrimination is strongest for the High-risk category and weakest for the Medium-risk category. The ROC curves are presented in Figure 3.

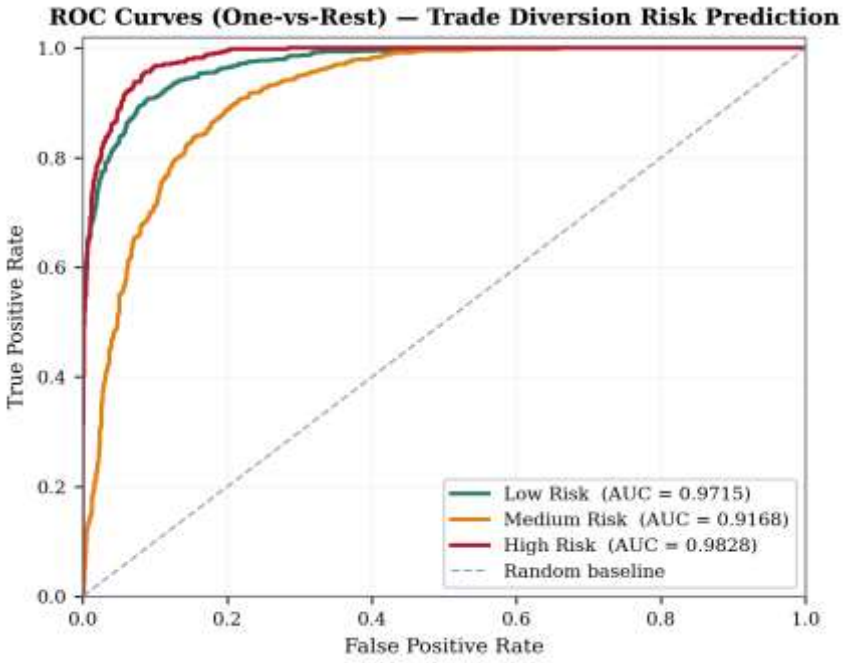

Figure 3. One-versus-rest ROC curves for the three risk categories.

Risk dimensions and feature representation

The engineered indicators can be ranked, and the ordering they generate is meaningful in supply-chain terms. The top three components of the specified risk-generating mechanism are Export Vulnerability, Diversion Exposure and Import Dependency. These three occupy the top three positions in all ten seeds, and Export Vulnerability ranks first in every one; the second and third positions alternate between Diversion Exposure and Import Dependency, so the top-three set is stable although the ordering within it is not. The feature-importance ranking is shown in Figure 4. The engineered indicators result in a stable and interpretable ordering. Export Vulnerability is the top-ranked indicator, since it reflects the interaction of export exposure and lack of flexibility, the profile of a firm that is unable to withstand a demand shock without harming its supply linkages. The markets that motivate the trade-diversion scenario are captured in Diversion Exposure. The upstream input channel has a significant contribution to the risk structure, as it carries the third-largest generating weight (0.20) and receives the third-highest rank from the model, which is in line with the emphasis that the Indian literature places on imported intermediates. Resilience Gap, the gap between awareness and flexibility, ranks fourth in all ten seeds. Trade Concentration has no weight in the generating mechanism and is ranked low in the importance ranking; its incremental effect is evaluated using the leave-one-out analysis below. The attribution results are interpreted primarily as evidence of ordering (Hooker et al., 2021; Molnar et al., 2022). Two checks support this interpretation of the ranking. Permutation importance, calculated in parallel with the impurity measure, agrees almost exactly with the ordering (Spearman's $\rho = 0.99$ over the eleven variables) , which corroborates that the ordering is not solely due to the impurity measure mentioned in Section 2.8. Leave-one-out ablation then examines how much the model's output changes when each indicator is omitted from the model's input. Removing Export Vulnerability or Import Dependency costs the Gradient Boosting model 0.066 and 0.068 macro-F1, respectively; removing Resilience Gap costs 0.034 and Diversion Exposure 0.027; and removing Trade Concentration changes performance by −0.0004, a change smaller than the seed-to-seed spread and positive in three of ten seeds. The ablation indicates a negligible incremental contribution from Trade Concentration when the other engineered indicators are retained in the model, which is consistent with the zero direct weight of Trade Concentration in Equation (6). The dependence diagnostics for the derived features are reported in detail in the supplementary material. A further consequence of the representation is practical. The five engineered indicators do not differ in performance from the full eleven-variable set for any learner tested, and all differences are within 0.0030 macro-F1, none of which is statistically distinguishable from zero. The five composite indicators thus constitute a compact representation whose predictive performance is statistically indistinguishable from that of the full set of eleven variables, reducing the dimensionality of the representation used for modelling and monitoring.

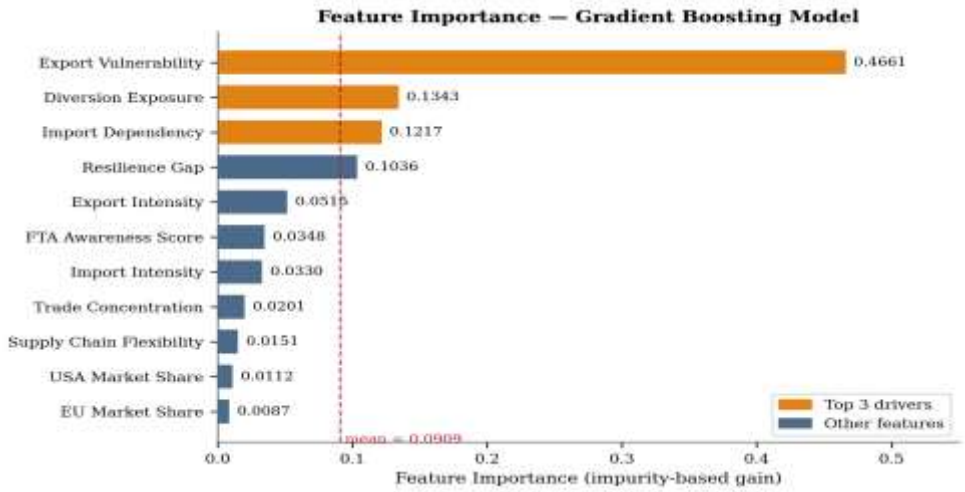

Figure 4. Feature importance in the Gradient Boosting model.

Feature engineering and learner capacity

The main technical finding of the study is the relationship between the construction of the features and the capacity of the learners, which is monotone over the four learners. The indicators were most useful for the Decision Tree and least useful for Gradient Boosting, suggesting that domain-informed indicators can make up for a lack of model capacity. The gain differs by a factor of more than eight, falling from +0.0528 macro-F1 for the Decision Tree to +0.0370 for Logistic Regression, +0.0186 for Random Forest and +0.0062 for Gradient Boosting, with non-overlapping confidence intervals between the simpler learners and the ensembles. The gains, confidence intervals, effect sizes and Holm-adjusted p-values for each learner are reported in Table 3.

Table 3. Effect of engineered indicators by learner (Δ macro-F1, raw versus raw + engineered, n = 10 seeds).

Learner	Raw only	Raw + engineered	Gain	95% CI	Cohen's dz	Holm p
Decision Tree	0.7342	0.7870	+0.0528	[+0.043, +0.063]	3.91	5.4e-6
Logistic Regression	0.7587	0.7956	+0.0370	[+0.033, +0.041]	6.72	4.2e-8
Random Forest	0.8191	0.8378	+0.0186	[+0.016, +0.021]	5.52	3.0e-7
Gradient Boosting	0.8364	0.8425	+0.0062	[+0.001, +0.011]	0.88	0.106

The gain is positive for all learners, but varies across models in terms of statistical support. The gain of Gradient Boosting (+0.0062 macro-F1) is not statistically significant after Holm adjustment (Holm p = 0.106), while the gains of Decision Tree, Logistic Regression and Random Forest are still statistically significant after Holm adjustment. The substantive reading of the trade-off between feature engineering and model capacity is therefore that they are two different routes to the same representational power. In the tested benchmark, explicit feature engineering partially compensates for lower learner capacity, but it is not a universal multiplier; a boosted ensemble can build interaction structure internally and gains least from being given it, whereas an additive model cannot and gains most. In the cases where the simpler models are more interpretable and/or less computationally intensive, the domain-informed indicators offer more incremental value; in the tested configurations, the high-capacity ensembles offer less incremental value from explicit feature construction. The two-tier pattern is supported by the specification tests: both lower-capacity learners gain more than both ensembles in each alternative specification, but the ordering within each pair is not stable, so the finding is a capacity gradient, not a four-way ranking.

Robustness and sensitivity analyses

The patterns above do not change across the tested specifications. The engineered-representation advantage is positive in all 28 robustness cells across seven alternative specifications of the risk structure varying the classification thresholds, the noise level and the destination-share supports, and is statistically distinguishable in 27 of these cells after adjusting for multiple comparisons. The only exception is at the highest noise setting, where no representation can recover structure destroyed by the noise. The advantage is positive in all 10 learner-configuration cells and statistically distinguishable in all 10 across the 5 alternative weight configurations. The capacity ordering is also stable: the two lower-capacity learners gain more than both ensembles in all seven specifications, and Gradient Boosting gains least in all seven, though the ordering within each pair is not stable, so the supported claim is a two-tier ordering, not a four-way ranking. The sector comparison serves as a negative-control check within the benchmark, as sector membership is independent of the risk structure. There is no statistically significant association between sector and risk category ($\chi^2 = 9.45$, df = 6, p = 0.150, Cramér's V = 0.022). The distribution of the sector is presented in Figure 5, while Figure 6 summarises the performance per class by the three risk categories.

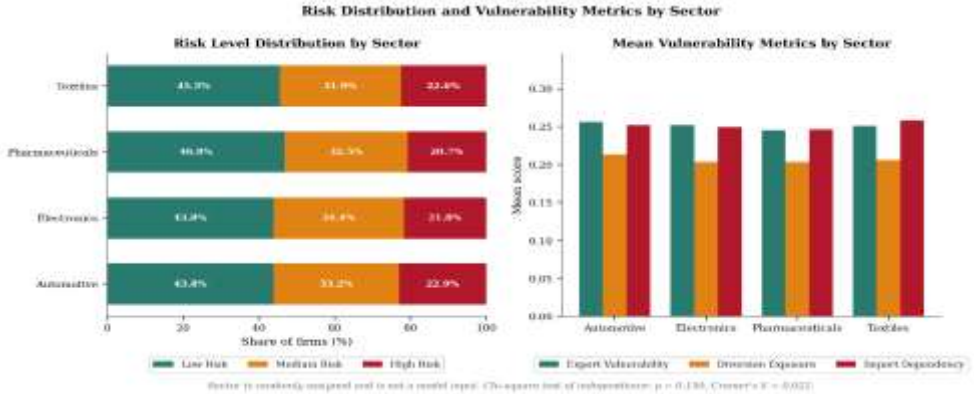

Figure 5. Risk distribution and mean engineered vulnerability indicators across sectors under the benchmark.



Figure 6. Precision, recall and F1-score by risk category.

DISCUSSIONS

The central finding is technical before it is managerial. The engineered representation results in higher macro-F1 for all learners tested, the five composite indicators perform as well as the full eleven-variable set, and the importance ordering is generally similar to the specified generating mechanism, with the zero-weight Trade Concentration still in the bottom half of the ordering. The representation thus makes the structure of the task more accessible to the learner, while at the same time keeping an interpretable structure based on the underlying constructs. This is an operational version of a longstanding debate in supply-chain management: vulnerability is not exposure, and measures that conflate them may misidentify which firms are at risk (Ambulkar et al., 2015; Pettit et al., 2010; Wagner & Bode, 2006). Export Vulnerability, which is a combined measure of export intensity and the inverse of operational flexibility, is exactly this type of measure and is first in all specifications tested. The second finding concerns the value of that representation and for whom it is greatest. The interaction between feature construction and learner capacity is the substantive result: a boosted ensemble approximates interaction structure internally and gains least from being handed it, while an additive learner cannot and gains most, so the two are partial substitutes rather than complements. The practical implication is that the interpretable representation is most beneficial for the simpler learners considered here, where the domain structure is not as easily constructed by the model. The construction of indicators also provides an interpretable representation that can be used for different model implementations. The upstream result deserves emphasis for supply-chain readers. Import Dependency was ranked third out of eleven variables, and was more important than all of the raw inputs within the benchmark, meaning that the input side was comparably important to the destination side within the benchmark. This is in line with the Indian evidence of imported intermediates (Ahsan, 2013; Arneja & Sharma, 2026; Goldberg et al., 2010) and indicates that contingency planning based on the concentration of buyers alone may fail to account for a material upstream exposure (Namdar et al., 2018). The sector check in Section 4.4 points in the same direction: the argument from the heterogeneous-firms literature (Bernard et al., 2012; Melitz, 2003) becomes a testable proposition rather than a result, because sector carries no classification signal in the benchmark.

Implications for Practice

For managers

Risk screening should consider exposure in conjunction with adaptability, not in isolation. A screening measure is more informative about vulnerability when it includes both exposure and the ability of the firm to redirect. The indicators are interpretable, and a firm's classification can be linked to specific exposure and capability dimensions, consistent with evidence that flexibility and visibility influence the way firms absorb trade and demand shocks (Brandon-Jones et al., 2014; Jüttner & Maklan, 2011).

For supply-chain analysts

The choice between investing in feature construction and investing in model sophistication can be made deliberately: the experiments suggest that feature construction is more beneficial for interpretable or lower-capacity learners and less beneficial for high-capacity ensembles. The experiments lead to two additional design points. First, a reduced set of five indicators lowers the dimensionality carried into modelling and monitoring. Second, since the indicators are based on common underlying variables, importance measures on them should be interpreted as evidence of ordering, not of magnitude. Attention should also be given to the upstream channel, as input dependence was a key risk dimension in the benchmark.

For policymakers and industry bodies

The framing is more granular than sector averages. The proposed indicators could be used in screening based on destination exposure, import dependence and adaptive capability, not just sector membership, where administrative trade and registration data are available at firm level. This offers a potential foundation for more specific identification of risk and trade adjustment assistance.

Limitations and Future Research

The study has several limitations. First, the evaluation is based on a controlled synthetic benchmark where the risk structure is specified rather than observed, and thus the results reflect the behaviour of the representation under controlled conditions, not predictive validity for realised firm outcomes. Second, there is no direct label of risk, as the categories are based on the Network Risk Score which is not based on actual outcomes like losing orders or discontinuing exports. Third, the data-generating process is assumed to be independent draws and does not replicate the correlation structure, temporal dynamics or sector effects that are present in operational data. Fourth, the risk-score weights are analyst-specified, although the overall results are consistent with the five alternative weights examined. Finally, hyperparameters were not tuned and probability calibration was not assessed, so the predicted probabilities should not be interpreted as calibrated risk estimates. Application to observed firm-level outcomes is required before operational deployment. Future research should test the indicators on firm-level trade or administrative data, compare them with independently observed events (e.g., export discontinuation or supplier disruption), incorporate realistic dependence structures, and add temporal validation. Temporal validation, in which indicators computed before a trade realignment are tested against outcomes recorded after it, would be an important operational test, and would also determine whether the capacity gradient observed here is maintained when exposure and capability are correlated as they are in practice.

CONCLUSIONS

This study proposed a structured representation of firm-level trade vulnerability that integrates market exposure with adaptive capability, and examined whether this representation is useful for risk classification under a given risk-generating mechanism. In the controlled synthetic benchmark, Gradient Boosting achieved an accuracy of 84.44 per cent, with cross-validated accuracy of 84.54 per cent ± 0.68 per cent and AUC above 0.91 in all categories, with errors clustered near the boundaries of the adjacent categories. Three findings go beyond the basic configuration of the model. First, the top three risk components are Export Vulnerability, Diversion Exposure and Import Dependency, with the upstream input channel comparable in importance to destination exposure. Second, as model capacity increases, the value of explicit indicator construction decreases, from +0.0528 macro-F1 for a decision tree to +0.0062 for gradient boosting, the latter not surviving multiplicity adjustment; domain feature engineering thus partially substitutes for learner capacity. Third, a compact five-indicator representation performed as well as the full eleven-variable set. The contribution is a domain-informed representation and a controlled way of testing the value of the contribution across learners. The findings serve as a foundation for expanding the indicators to firm-level trade data and other risk-screening uses.

Author Contributions: Conceptualization, N.P., R.R., S.R. and C.S.D.; Methodology, N.P.; Software, N.P.; Validation, S.R.; Formal Analysis, N.P., R.R., S.R. and C.S.D.; Investigation, N.P.; Resources, N.P.; Data Curation, N.P.; Writing – Original Draft Preparation, N.P., R.R., S.R. and C.S.D.; Writing – Review & Editing, N.P., R.R., S.R. and C.S.D.; Visualization, N.P.; Supervision, S.R.; Project Administration, N.P.; Funding Acquisition, N.P., R.R., S.R. and C.S.D. Authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Ethical review and approval were waived for this study because the research does not involve vulnerable groups or sensitive issues.

Funding: The authors received no direct funding for this research.

Acknowledgements: The authors have no acknowledgements to declare.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to restrictions.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process: During the preparation of this work, the author(s) used Grammarly for proofreading and spell checking since the Authors are not native speakers. All intellectual content, analysis, and interpretations were produced solely by the authors. After using this AI tool/service, the author(s) reviewed and edited the content as needed, taking full responsibility for the publication's content.

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

Ahsan, R. N. (2013). Input tariffs, speed of contract enforcement, and the productivity of firms in India. *Journal of International Economics, 90*(1), 181–192. https://doi.org/10.1016/j.jinteco.2012.11.006

Ambulkar, S., Blackhurst, J., & Grawe, S. (2015). Firm's resilience to supply chain disruptions: Scale development and empirical examination. *Journal of Operations Management, 33–34*(1), 111–122. https://doi.org/10.1016/j.jom.2014.11.002

Arneja, N., & Sharma, C. (2026). Imported intermediates and firm performance: A difference-in-differences analysis of matched Indian manufacturing firms. *Technological Forecasting and Social Change, 226*, Article 124571. https://doi.org/10.1016/j.techfore.2026.124571

Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J., & Vanthienen, J. (2003). Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of the Operational Research Society, 54*(6), 627–635. https://doi.org/10.1057/palgrave.jors.2601545

Baldwin, R., & Freeman, R. (2022). Risks and global supply chains: What we know and what we need to know. *Annual Review of Economics, 14*(1), 153–180. https://doi.org/10.1146/annurev-economics-051420-113737

Banerjee, S., Karna, A., & Sharma, S. (2022). Imprinting effects of exposure to the Indian independence movement on export intensity of firms. *Journal of Business Research, 149*, 548–557. https://doi.org/10.1016/j.jbusres.2022.05.033

Bernard, A. B., Jensen, J. B., Redding, S. J., & Schott, P. K. (2007). Firms in international trade. *Journal of Economic Perspectives, 21*(3), 105–130. https://doi.org/10.1257/jep.21.3.105

Bernard, A. B., Jensen, J. B., Redding, S. J., & Schott, P. K. (2012). The empirics of firm heterogeneity and international trade. *Annual Review of Economics, 4*(1), 283–313. https://doi.org/10.1146/annurev-economics-080511-110928

Bhattacharjee, J., Puri, L., Singh, R., Dubey, V., & Pardeshi, D. (2026). Factors shaping Indian Premier League team fan loyalty: A social network analysis approach. *Managing Sport and Leisure*. Advance online publication. https://doi.org/10.1080/23750472.2026.2663923

Brandon-Jones, E., Squire, B., Autry, C. W., & Petersen, K. J. (2014). A contingent resource-based perspective of supply chain resilience and robustness. *Journal of Supply Chain Management, 50*(3), 55–73. https://doi.org/10.1111/jscm.12050

Caldara, D., Iacoviello, M., Molligo, P., Prestipino, A., & Raffo, A. (2020). The economic effects of trade policy uncertainty. *Journal of Monetary Economics, 109*, 38–59. https://doi.org/10.1016/j.jmoneco.2019.11.002

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). Association for Computing Machinery. https://doi.org/10.1145/2939672.2939785

Chowdhury, M. M. H., & Quaddus, M. (2017). Supply chain resilience: Conceptualization and scale development using dynamic capability theory. *International Journal of Production Economics, 188*, 185–204. https://doi.org/10.1016/j.ijpe.2017.03.020

Craighead, C. W., Blackhurst, J., Rungtusanatham, M. J., & Handfield, R. B. (2007). The severity of supply chain disruptions: Design characteristics and mitigation capabilities. *Decision Sciences, 38*(1), 131–156. https://doi.org/10.1111/j.1540-5915.2007.00151.x

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., & Pavcnik, N. (2016). Prices, markups, and trade reform. *Econometrica, 84*(2), 445–510. https://doi.org/10.3982/ECTA11042

Dev, D., Deb, S., Singh, R., & Puri, L. (2022). Assessing relative weight of determinants of investment in index scheme of mutual funds. *Indian Journal of Finance and Banking, 11*(1), 29–37. https://doi.org/10.46281/ijfb.v11i1.1847

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM, 55*(10), 78–87. https://doi.org/10.1145/2347736.2347755

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics, 29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Fugazza, M., & Nicita, A. (2013). The direct and relative effects of preferential market access. *Journal of International Economics, 89*(2), 357–368. https://doi.org/10.1016/j.jinteco.2012.09.001

Ghosh, M., & Sinha Roy, S. (2018). Foreign direct investment, firm heterogeneity, and exports: An analysis of Indian manufacturing. *Asian Development Review, 35*(1), 27–51. https://doi.org/10.1162/adev_a_00104

Goldberg, P. K., Khandelwal, A. K., Pavcnik, N., & Topalova, P. (2010). Imported intermediate inputs and domestic product growth: Evidence from India. *The Quarterly Journal of Economics, 125*(4), 1727–1767. https://doi.org/10.1162/qjec.2010.125.4.1727

Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on typical tabular data? *Advances in Neural Information Processing Systems, 35*, 507–520. https://doi.org/10.52202/068431-0037

Grossman, G. M., Helpman, E., & Lhuillier, H. (2023). Supply chain resilience: Should policy promote international diversification or reshoring? *Journal of Political Economy, 131*(12), 3462–3496. https://doi.org/10.1086/725173

Handley, K., & Limão, N. (2017). Policy uncertainty, trade, and welfare: Theory and evidence for China and the United States. *American Economic Review, 107*(9), 2731–2783. https://doi.org/10.1257/aer.20141419

Heaton, J. (2016). An empirical analysis of feature engineering for predictive modeling. In *SoutheastCon 2016* (pp. 1–6). IEEE. https://doi.org/10.1109/SECON.2016.7506650

Hooker, G., Mentch, L., & Zhou, S. (2021). Unrestricted permutation forces extrapolation: Variable importance requires at least one more model, or there is no free variable importance. *Statistics and Computing, 31*(6), Article 82. https://doi.org/10.1007/s11222-021-10057-z

Hosseini, S., Ivanov, D., & Dolgui, A. (2019). Review of quantitative methods for supply chain resilience analysis. *Transportation Research Part E: Logistics and Transportation Review, 125*, 285–307. https://doi.org/10.1016/j.tre.2019.03.001

Ivanov, D., & Dolgui, A. (2020). Viability of intertwined supply networks: Extending the supply chain resilience angles towards survivability. A position paper motivated by COVID-19 outbreak. *International Journal of Production Research, 58*(10), 2904–2915. https://doi.org/10.1080/00207543.2020.1750727

Jüttner, U., & Maklan, S. (2011). Supply chain resilience in the global financial crisis: An empirical study. *Supply Chain Management: An International Journal, 16*(4), 246–259. https://doi.org/10.1108/13598541111139062

Kashiwagi, Y., Todo, Y., & Matous, P. (2021). Propagation of economic shocks through global supply chains—Evidence from Hurricane Sandy. *Review of International Economics, 29*(5), 1186–1220. https://doi.org/10.1111/roie.12541

Kuhn, M., & Johnson, K. (2019). *Feature engineering and selection: A practical approach for predictive models*. Chapman and Hall/CRC. https://doi.org/10.1201/9781315108230

Lessmann, S., Baesens, B., Seow, H.-V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research, 247*(1), 124–136. https://doi.org/10.1016/j.ejor.2015.05.030

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica, 71*(6), 1695–1725. https://doi.org/10.1111/1468-0262.00467

Molnar, C., König, G., Herbinger, J., Freiesleben, T., Dandl, S., Scholbeck, C. A., Casalicchio, G., Grosse-Wentrup, M., & Bischl, B. (2022). General pitfalls of model-agnostic interpretation methods for machine learning models. In A. Holzinger, R. Goebel, R. Fong, T. Moon, K.-R. Müller, & W. Samek (Eds.), *xxAI – Beyond explainable AI* (pp. 39–68). Springer. https://doi.org/10.1007/978-3-031-04083-2_4

Namdar, J., Li, X., Sawhney, R., & Pradhan, N. (2018). Supply chain resilience for single and multiple sourcing in the presence of disruption risks. *International Journal of Production Research, 56*(6), 2339–2360. https://doi.org/10.1080/00207543.2017.1370149

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*(85), 2825–2830. https://www.jmlr.org/papers/v12/pedregosa11a.html

Pettit, T. J., Fiksel, J., & Croxton, K. L. (2010). Ensuring supply chain resilience: Development of a conceptual framework. *Journal of Business Logistics, 31*(1), 1–21. https://doi.org/10.1002/j.2158-1592.2010.tb00125.x

Pierce, J. R., & Schott, P. K. (2016). The surprisingly swift decline of US manufacturing employment. *American Economic Review, 106*(7), 1632–1662. https://doi.org/10.1257/aer.20131578

Reddy, K., & Sasidharan, S. (2021). Financial constraints and global value chain participation: Firm-level evidence from India. *The Journal of International Trade & Economic Development, 30*(5), 739–765. https://doi.org/10.1080/09638199.2021.1900343

Reddy, K., & Sasidharan, S. (2024). Global value chains, productivity and markup: Evidence from India. *International Review of Economics & Finance, 89*, 250–271. https://doi.org/10.1016/j.iref.2023.10.005

Reddy, K., Sasidharan, S., & Thangavelu, S. M. (2025). Does economic policy uncertainty impact global value chain participation of Indian firms? The role of financial constraints. *Economic Modelling, 152*, Article 107243. https://doi.org/10.1016/j.econmod.2025.107243

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence, 1*(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

Shwartz-Ziv, R., & Armon, A. (2022). Tabular data: Deep learning is not all you need. *Information Fusion, 81*, 84–90. https://doi.org/10.1016/j.inffus.2021.11.011

Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics, 8*(1), Article 25. https://doi.org/10.1186/1471-2105-8-25

Todo, Y., Nakajima, K., & Matous, P. (2015). How do supply chain networks affect the resilience of firms to natural disasters? Evidence from the Great East Japan Earthquake. *Journal of Regional Science, 55*(2), 209–229. https://doi.org/10.1111/jors.12119

Topalova, P., & Khandelwal, A. (2011). Trade liberalization and firm productivity: The case of India. *The Review of Economics and Statistics, 93*(3), 995–1009. https://doi.org/10.1162/rest_a_00095

Viner, J. (2014). *The customs union issue* (P. Oslington, Ed.). Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199756124.001.0001 (Original work published 1950)

Wagner, S. M., & Bode, C. (2006). An empirical investigation into supply chain vulnerability. *Journal of Purchasing and Supply Management, 12*(6), 301–312. https://doi.org/10.1016/j.pursup.2007.01.004

Publisher's Note: ACSE stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2026 by the authors. Licensee ACSE, USA. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

American International Journal of Sciences and Engineering Research (*E-ISSN 2641-0311 P-ISSN 2641-0303)* by ACSE is licensed under a Creative Commons Attribution 4.0 International License.