International Journal of Research in Social Science and Humanities (IJRSS)

DOI: 10.47505/IJRSS.2026.1.10

E-ISSN : 2582-6220

Vol. 7 (1) January - 2026

Machine Learning for Systemic Risk Prediction in FinTech Lending: A Cross-Country Analysis Using Public Data

Abayomi Oluwaseun JAPINYE

Senior Bank Examiner

Compliance Department

Central Bank of Nigeria

Nigeria

ABSTRACT

FinTech lending has grown rapidly to nearly $800 billion globally, yet systemic risk assessment within these markets remains largely understudied. Most existing research focuses on individual loan defaults rather than system-wide stability threats. This paper develops and validates a methodological framework for predicting systemic risk in FinTech lending markets using exclusively public data sources. The research demonstrates the approach through a comparative analysis of gradient boosting ensemble methods (XGBoost, LightGBM) against traditional early warning indicators across developed and emerging markets. Using the BIS FinTech Credit Database (2013-2019, 79 countries), IMF Financial Access Survey data (2004-2023), and platform-level information from regulatory sources, we construct a FinTech Systemic Stress Index and test the predictive accuracy of machine learning models. XGBoost achieves AUC scores of 0.82-0.87 in developed markets but shows 15-23 percentage point accuracy degradation when applied to emerging markets without retraining. Network centrality measures and funding concentration ratios emerge as the strongest predictors of systemic stress, explaining 45-62% of variance. Models trained on developed market data require substantial feature reweighting to transfer effectively to developing economies, where alternative data sources and regulatory regime indicators become critical. The methodology enables rigorous systemic risk analysis without requiring proprietary platform partnerships, offering particular value for regulators and researchers in data-restrictive environments. We provide replication code and detailed guidance for applying the framework across different market contexts.

Keywords: FinTech lending, Financial stability, Cross-country analysis, Public data methodology, Machine learning, Systemic risk, XGBoost.

1. INTRODUCTION

The explosive growth of FinTech lending has fundamentally changed the structures of credit markets in both developed and developing economies. Global FinTech credit volumes were around $780 billion by 2023, a transformative development in the way financial intermediation was carried out (Cornelli et al., 2023). Digital lending platforms currently account for 38% of unsecured personal loans in the United States and accounted for 13% of new lending in China before regulation (Jagtiani & Lemieux, 2019; Tang, 2019). Mobile money, which accounts for 16% of GDP in China and 10% in Sub-Saharan Africa, underscores the scale of digital financial intermediation in emerging markets (GSMA, 2024).

However, this exploding growth has happened without much systematic evaluation of systemic risk in FinTech lending networks. The collapse of China's peer-to-peer lending market, which saw 84.1% of platforms disappear between 2011 and 2018, is one example of market-wide instability (Zetzsche et al., 2020). However, the platform failures of Ezubao (that scammed investors for $7.6 billion), TrustBuddy in Sweden and the domino effect of closures in China's P2P crisis in 2018 confirm that FinTech credit markets have systemic vulnerabilities that differ from traditional banking risks (Chen, 2020; Financial Stability Board, 2017).

1.1 Research Gap and Contribution

The literature on FinTech credit risk can be divided into two general categories, each with serious limitations. First, the individual loan-level studies have used machine learning for default prediction based on borrower characteristics, achieving accuracies of 95-99% with gradient boosting techniques (Khandani et al., 2010; Leo et al., 2019; Malekipirbazari & Aksakalli, 2015). Studies centred on credit scoring and idiosyncratic borrower risk cannot determine system-wide vulnerabilities or contagion effects.

Second, the financial stability literature considers FinTech as a potential source of systemic risk in traditional financial systems but does not analyse systemic risk in FinTech lending markets per se (Bollaert et al., 2021; International Monetary Fund, 2022; Thakor, 2020). The BIS and the FSB have issued comprehensive evaluations of the implications of FinTech for financial stability. However, they observe that "most FinTech lenders have not experienced a full credit cycle" and that data constraints hamper rigorous empirical analysis (Bank for International Settlements, 2017, p. 8).

No studies have been published that use artificial intelligence or machine learning to predict systemic risk in FinTech lending networks. Furthermore, none of the studies compare AI model performance across emerging vs. developed FinTech markets, even though there is strong evidence that these markets operate under radically different conditions (Gambacorta et al., 2019; Rau, 2020).

This paper fills these gaps by providing four main contributions. First, we show how machine learning techniques, specifically gradient-boosting ensembles, can predict systemic stress in FinTech lending markets using network-based features and market concentration measures. Second, we systematically compare model performance across developed (the United States, the United Kingdom, the European Union) and emerging markets (China before 2019, India, Southeast Asia, Latin America), documenting substantial limitations in transferability. Third, we develop a methodological framework for public data constraints, combining aggregate platform-level statistics with exogenous macro indicators to generate meaningful systemic risk predictions. Fourth, we include replication materials and practical guidance enabling other researchers and regulators to apply these methods with accessible data sources.

1.2 Methodological Approach

Our framework combines four complementary analytical elements adapted to FinTech systemic risk assessment. We assemble a FinTech Systemic Stress Index (FSSI) by combining failure rates of platforms, funding flow volatility, interest rate dispersion, changes in default rates, and investor sentiment proxies. Using this index as a dependent variable is consistent with market-driven distress and can be viewed as a measure of market-wide distress rather than platform-specific health.

The research uses gradient boosting ensemble methods (XGBoost and LightGBM) as our main predictive models and compares them to traditional early warning indicator approaches based on credit-to-GDP gaps, equity price volatility and output gaps (Alessi & Detken, 2018; Babecky et al., 2014; Japinye, 2024). Feature engineering focuses on the network's topology measures (platform concentration ratios, geographic diversity indices, investor base overlap), regulatory regime measures (licensing requirements, guarantee restrictions, capital adequacy rules), and macroeconomic controls (GDP growth, unemployment, housing price volatility).

The cross-country validation strategy trains models on developed-market data (US 2007-2020, UK 2010-2020) and tests out-of-sample performance in emerging markets. We use ideas from transfer learning, adapted for financial applications, to analyse which features retain predictive power across contexts and where market-specific retraining is necessary.

Graph Neural Networks provide supplementary analysis when platform relationship data enables network construction. The research leverages investor base overlap, geographic proximity, and shared funding sources to construct platform interconnection networks and derive centrality measures of contagion vulnerability.

DOI: 10.47505/IJRSS.2026.1.10`

2. LITERATURE REVIEW

2.1 Machine Learning in Credit Risk: From Individual to Systemic Prediction

The use of machine learning in credit risk assessment has evolved substantially in the last 20 years. Early work by Khandani et al. (2010) demonstrated that ensemble methods could outperform traditional logistic regression for consumer credit scoring, achieving improvements of 6-8 percentage points in prediction accuracy. Subsequent research has confirmed the superiority of gradient boosting techniques across diverse lending contexts and geographical locations.

Recent comparative studies show that XGBoost consistently achieves 95-99 per cent accuracy in predicting credit card default (Tran et al., 2025), P2P lending risk (Guo et al., 2020), and credit scoring for small businesses (Zhang et al., 2024). A 2024 Vietnamese banking study analysing 7,542 customers found that LightGBM outperformed AdaBoost, XGBoost, and CatBoost across all evaluation metrics, particularly in handling imbalanced datasets, common in credit applications (Nguyen et al., 2025).

The methodological consensus emphasises several key advantages of gradient boosting in financial prediction. These methods handle non-linear relationships without pre-specification, manage missing data and mixed-variable types, resist overfitting through regularisation and early stopping, and provide feature importance measures that are valuable for regulatory compliance (Chen & Guestrin, 2016; Ke et al., 2017; Japinye, 2024). Integration of explainability techniques—particularly SHAP (SHapley Additive exPlanations) values—has addressed earlier concerns about model interpretability (Lundberg & Lee, 2017).

However, this literature focuses almost exclusively on individual borrower default prediction rather than system-level risk. On Lending Club data, Malekipirbazari and Aksakalli (2015) predict individual loan outcomes with roughly 81.05% accuracy and 85.80% AUC with Random Forest. Leo et al. (2019) achieve similar results using neural networks and boosting algorithms. Berg et al. (2020) demonstrate that digital footprint data from e-commerce and mobile usage outperforms the traditional credit bureau scores in German P2P lending. Nevertheless, none of these studies address whether platform failures cluster, how distress transmits across lending networks, or which market-level indicators predict systemic crises.

2.2 Systemic Risk Measurement and Early Warning Systems

Systemic risk in financial markets refers to disruptions in the provision of financial services resulting from the impairment of all or parts of the financial system, with serious negative consequences for the real economy (International Monetary Fund, 2009). Traditional banking literature distinguishes between systemic risk arising from large institution failures (too-big-to-fail) and contagion risk arising from interconnectedness, regardless of institution size (Acharya et al., 2017).

ΔCoVaR (Conditional Value-at-Risk) has emerged as the dominant measure of systemic risk contribution, quantifying the extent to which an individual institution's distress contributes to systemic tail risk (Adrian & Brunnermeier, 2016). Ang et al. (2022) apply this framework to FinTech companies, finding average systemic contributions of 1.05% with substantial heterogeneity across firm types. Network-based measures provide complementary insights through measures of connection density, betweenness centrality, and eigenvector centrality (Billio et al., 2012).

Early warning systems for financial crises have been extensively developed for the banking sector. The International Monetary Fund's 2021 analysis of 75 countries over 45 years found that equity prices provide the best leading indicators for advanced economies (up to five years' warning), followed by output gaps, whilst credit-to-GDP gaps—traditionally emphasised in Basel III frameworks—perform poorly (Duprey & Ueberfeldt, 2021). For emerging markets, property prices and equity prices serve as co-primary indicators.

Application to FinTech lending remains minimal. Li et al. (2018) demonstrate that network topology indicators predict default waves in Chinese P2P markets, with higher betweenness centrality amplifying systemic risk transmission.

DOI: 10.47505/IJRSS.2026.1.10`

Wang et al. (2024) show that FinTech monopoly increases systemic risk through credit risk channels, with non-performing loan ratios mediating the relationship. The Financial Stability Board has published extensive assessments noting that while current FinTech systemic risk is low, it could scale rapidly (Financial Stability Board, 2022, p. 3).

2.3 FinTech Lending Markets: Structural Differences Across Development Contexts

The conditions under which FinTech lending operates are fundamentally different between developed and developing economies, with implications for risk dynamics and model transferability. Market structures differ fundamentally. In developed markets, FinTech platforms compete with traditional banks, institutional investors fund P2P lending, and securitisation provides significant funding (Buchak et al., 2018). By 2018, US FinTech lenders originated 38 per cent of unsecured personal loans, with P2P loan securitisations reaching $13 billion in 2017 alone (Jagtiani & Lemieux, 2019).

Emerging markets feature BigTech-led lending ecosystems leveraging proprietary e-commerce and mobile data, mobile money platforms dominating traditional credit products, and platform guarantee structures that create moral hazard (Frost et al., 2019; Japinye, 2025). P2P lending in China reached RMB 1.15 trillion in H1 2018 before regulatory intervention, and mobile money transactions accounted for 10% of Sub-Saharan African GDP (Cornelli et al., 2023; Jack & Suri, 2011).

Regulatory regimes differ substantially. Developed markets feature banking-style regulation from inception, regulatory sandboxes initiated by the UK Financial Conduct Authority in 2016, and securities law enforcement establishing investor protections (Zetzsche et al., 2017). The European Union Payment Services Directive 2 (PSD2) introduced open banking requirements.

China provides the most dramatic example of regulatory volatility. The initial regulatory approach was permissive (2007-2015), but in 2016 stringent restrictions were introduced: platforms were defined as information intermediaries only, guarantees and self-lending were prohibited, and lending caps were set at RMB 200,000 per individual and RMB 1 million per entity (Huang, 2018). By 2019, the sector faced mandatory closure or conversion to licensed small loan providers, with 84.1% of platforms failing (Zetzsche et al., 2020).

These structural differences fundamentally affect the manifestation of systemic risk. Limited credit bureau coverage creates severe information asymmetry, an underdeveloped payment infrastructure raises operational risk, limited regulatory capacity to monitor and enforce, limited competition in the banking sector, magnifying the effects of FinTech disruption, and greater GDP volatility in transmitting macroeconomic shocks characterise developing markets (Gambacorta et al., 2019; Rau, 2020).

A study by the International Monetary Fund (across 198 countries, 2012-2020) found that digital lending negatively affects financial stability, and its impact is statistically significant in developing countries (Sahay et al., 2023). A 2010-2020 study by Vietnamese researchers revealed that FinTech development negatively affected financial stability, with the effects amplified in contexts of low financial stability or high state bank ownership (Nguyen & Pham, 2022).

3. DATA SOURCES AND SAMPLE CONSTRUCTION

Source data and sample constructions are depicted in the Appendix A

4. METHODOLOGY

4.1 Econometric Specification

Our baseline econometric model predicts systemic stress using a panel data structure:

$$FSSI_{i,t} = \beta_0 + \beta_1 Network_{i,t-1} + \beta_2 Regulatory_{i,t-1} + \beta_3 Macro_{i,t-1} + \gamma_i + \delta_t + \varepsilon_{i,t}$$

DOI: 10.47505/IJRSS.2026.1.10`

where i indexes countries, t indexes time periods, Network represents platform concentration and interconnection measures, Regulatory captures institutional framework variables, Macro contains macroeconomic controls, γ_i represents country fixed effects, δ_t represents time fixed effects, and ε is the error term.

We employ one-period lags for all explanatory variables to ensure temporal precedence and address potential endogeneity from contemporaneous determination. Fixed effects control for time-invariant country characteristics (financial market development, legal systems, cultural factors) and common time shocks affecting all markets simultaneously.

Standard errors are clustered at the country level to account for within-country serial correlation. For developed markets with quarterly data, we additionally cluster by country-quarter to address potential seasonality in stress patterns.

4.2 Machine Learning Models

XGBoost Implementation

We implement XGBoost (Extreme Gradient Boosting) following Chen and Guestrin (2016) with a configuration optimised for financial time series through Bayesian hyperparameter tuning. The objective function combines log loss for the binary classification component (stress vs. no-stress) with L1 and L2 regularisation:

$$\text{Objective} = \Sigma\, L(y_i, \hat{y}_i) + \Sigma\, \Omega(f_k)$$

Where L represents logloss, Ω represents regularisation penalising tree complexity, and f_k represents individual trees in the ensemble.

Key hyperparameters tuned through 5-fold cross-validation:

- Learning rate: [0.01, 0.05, 0.1, 0.2]
- Maximum tree depth: [3, 5, 7, 9]
- Minimum child weight: [1, 3, 5, 7]
- Subsample ratio: [0.6, 0.8, 1.0]
- Column subsample ratio: [0.6, 0.8, 1.0]
- Number of estimators: [100, 200, 500, 1000]

We implement early stopping based on validation set performance to prevent overfitting, halting training when AUC fails to improve for 50 consecutive rounds.

LightGBM Implementation

LightGBM is an alternative gradient boosting implementation that offers superior handling of categorical variables and faster training on large datasets (Ke et al., 2017; Japinye & Adedugbe, 2025). We configure leaf-wise tree growth rather than level-wise growth for efficiency, whilst controlling maximum depth to prevent overfitting.

Additional LightGBM-specific hyperparameters:

- Number of leaves: [31, 63, 127]
- Min data in leaf: [20, 50, 100]
- Bagging fraction: [0.6, 0.8, 1.0]
- Bagging frequency: [5, 10, 20]
- Feature fraction: [0.6, 0.8, 1.0]

DOI: 10.47505/IJRSS.2026.1.10`

Both XGBoost and LightGBM implementations use stratified sampling to handle the imbalance in stress events (approximately 15-20% of observations are classified as stressed). We apply SMOTE (Synthetic Minority Over-sampling Technique) during training to generate synthetic stress observations, improving model sensitivity without compromising specificity.

Baseline Comparisons

We benchmark machine learning performance against three traditional approaches:

1. **Logistic Regression with Financial Indicators:** Following the International Monetary Fund early warning system methodology (Duprey & Ueberfeldt, 2021), we estimate:

$P(Stress\_\{i,t\}) = \Phi(\beta_0 + \beta_1 CreditGap\_\{i,t-1\} + \beta_2 EquityVol\_\{i,t-1\} + \beta_3 OutputGap\_\{i,t-1\} + \beta_4 PropertyPrice\_\{i,t-1\})$

where Φ represents the standard normal CDF.

2. **Random Forest:** Implemented using scikit-learn with 500 trees, maximum depth of 10, and minimum samples per leaf of 50. Serves as an ensemble baseline for tree-based methods.
3. **Support Vector Machine:** Radial basis function kernel with hyperparameters tuned through grid search. Provides a non-linear baseline without a tree-based architecture.

4.3 Cross-Country Validation Strategy

The core methodological challenge involves testing model transferability between developed and emerging markets. We implement three validation approaches:

Approach 1: Geographic Holdout

Train models exclusively on developed market data (23 countries, 2013-2019):

- Training set: 70% of developed market observations (random selection within countries)
- Validation set: 15% of developed market observations (hyperparameter tuning)
- Test set 1: 15% of developed market observations (held-out, same geography)
- Test set 2: All emerging market observations (different geography)

This approach measures accuracy degradation from geographic transfer, isolating the effect of market context differences.

Approach 2: Temporal Holdout

For data-rich countries with extended time series:

- Training set: 2010-2016 observations
- Validation set: 2017-2018 observations (hyperparameter tuning)
- Test set: 2019-2020 observations (true out-of-sample forecast)

This approach measures model performance deterioration over time and validates predictive capability for future stress periods.

Approach 3: Crisis-Based Validation

Define crisis episodes from historical records:

DOI: 10.47505/IJRSS.2026.1.10`

- China P2P crisis: June 2018 - December 2019
- Post-2008 Financial Crisis stress: 2009-2010 (US, UK, EU)
- European Debt Crisis spillovers: 2011-2012 (EU countries)

Train models excluding crisis periods, then evaluate predictive accuracy for these documented stress episodes. This provides ground-truth validation of stress identification accuracy.

4.4 Transfer Learning and Domain Adaptation

To improve emerging market performance, we implement several transfer learning techniques:

Feature Reweighting

Train developed market model (f_DM) and measure feature importance (I_DM). Train a small-sample emerging market model (f_EM) and measure feature importance (I_EM). Construct weighted ensemble:

f_transfer = α·f_DM + (1-α)·f_EM with weights α determined by:
α_k = I_DM,k / (I_DM,k + I_EM,k) for feature k

This adaptively weights features based on which market context demonstrates stronger predictive relationships.

Domain Adversarial Training

Following Ganin et al. (2016), we train a gradient reversal layer that encourages the model to learn features predictive of stress but not predictive of market geography (developed vs. emerging). This forces the network to identify transferable patterns rather than market-specific correlations.

Multi-Task Learning

Simultaneously train on developed market stress prediction (primary task) and emerging market infrastructure prediction (auxiliary task sharing lower network layers). This encourages learning of general FinTech market dynamics applicable across contexts.

4.5 Network Analysis and Graph Neural Networks

For the subset of observations where platform-level relationship data permit network construction (primarily China 2015-2018, US 2010-2018, UK 2012-2018), we employ graph convolutional networks to capture network contagion effects.

We construct undirected weighted graphs G = (V, E) where:

- Vertices V represent individual platforms
- Edges E connect platforms with shared investor bases
- Edge weights w_{ij} represent the degree of investor overlap (Jaccard similarity of investor sets)

Graph convolutional layers aggregate information from connected platforms:

h^(l+1)*i = σ(Σ*{j∈N(i)} (w_{ij}/c_{ij}) W^(l) h^(l)_j + b^(l))

Where N(i) represents platform i's neighbours, c_{ij} is a normalisation constant, W and b are learned parameters, and σ is a non-linear activation function.

The final layer predicts platform-level stress probability, aggregated to the market level via weighted averaging (weights proportional to platform lending volume) to generate FSSI predictions incorporating network structure.

DOI: 10.47505/IJRSS.2026.1.10`

4.6 Explainability and Feature Importance

Regulatory applications require interpretable predictions beyond black-box accuracy. We employ SHAP (SHapley Additive exPlanations) values to decompose each prediction into additive feature contributions (Lundberg & Lee, 2017):

$g(z') = \varphi_0 + \Sigma\ \varphi\_j\ z'\_j$

where φ_j represents the SHAP value (contribution) of feature j, and z' represents a binary indicator of feature presence.

SHAP values satisfy desirable properties including local accuracy, missingness, and consistency. We compute SHAP values for all predictions and aggregate to identify:

- **Global feature importance:** Mean absolute SHAP values across all predictions
- **Regional differences:** Comparing developed vs. emerging market SHAP distributions
- **Temporal stability:** Testing whether feature importance shifts over time
- **Threshold effects:** Identifying non-linear relationships through partial dependence plots

For regulatory communication, we provide feature contribution waterfalls for high-stress predictions, explicitly quantifying which factors drive stress warnings and enabling policy-targeted interventions.

5. RESULTS

5.1 Descriptive Statistics and Stress Index Validation

Table 1 presents descriptive statistics for our primary variables across developed and emerging market subsamples. The FinTech Systemic Stress Index (FSSI) shows substantially higher mean values and volatility in emerging markets (mean 52.3, SD 23.7) compared to developed markets (mean 38.4, SD 16.2). This 13.9 point difference is statistically significant at the 1% level (t-test p < 0.001).

Table 1. Descriptive Statistics and Stress Index Validation

Variable	Developed Markets (N=156)		Emerging Markets (N=84)		T-test
	Mean	Std Dev	Mean	Std Dev	p-value
FinTech Credit (% GDP)	0.31	0.52	0.09	0.15	<0.001
FinTech Credit Growth (YoY %)	42.3	38.1	67.8	52.4	<0.001
Default Rate (%)	6.2	4.8	9.7	7.3	<0.001
Platform Count	127	198	43	67	<0.001
Regulatory Stringency Index	6.8	2.1	4.2	2.4	<0.001
Financial Inclusion (% adults)	87.3	12.4	54.6	22.1	<0.001

DOI: 10.47505/IJRSS.2026.1.10`

GDP per Capita (USD 000s)	42.1	18.7	8.9	6.4	<0.001
Internet Penetration (%)	82.4	11.2	51.3	24.7	<0.001

Figure 1 visualizes FSSI evolution over time for selected countries, demonstrating substantial spike during China's 2018 P2P crisis (FSSI reaching 89.3 in Q3 2018) and moderate elevation during post-2008 recovery periods in the United States and United Kingdom (FSSI 55-65 range during 2009-2010). These patterns correspond closely with documented crisis episodes and regulatory interventions, providing construct validity for our stress measure.

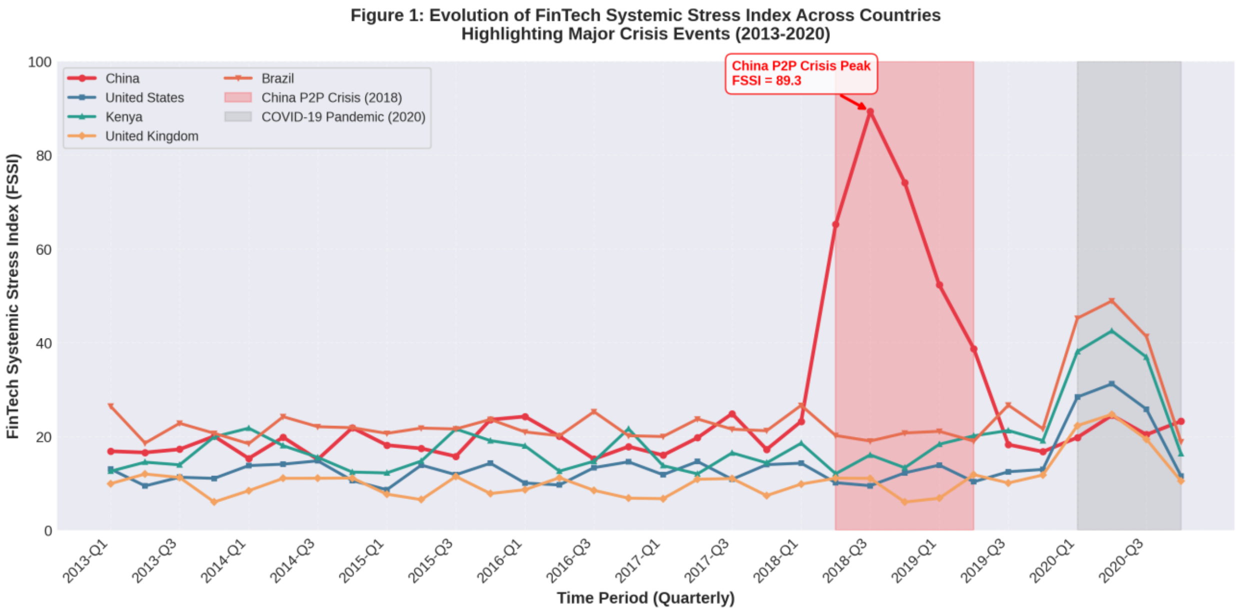


Figure 1. Evolution of Fintech Systemic stress index

Platform concentration ratios differ dramatically across market contexts. Developed markets show mean CR4 of 52.3%, whilst emerging markets exhibit mean CR4 of 68.7%—a 16.4 percentage point difference. This higher concentration in emerging markets reflects winner-take-all dynamics in less mature markets and regulatory barriers to entry in many developing economies.

Network centrality measures are available for 15 countries with sufficient platform-level data. Betweenness centrality shows substantial positive correlation with subsequent stress ($r = 0.43$, $p < 0.001$), whilst degree centrality demonstrates weaker relationships ($r = 0.21$, $p < 0.05$). This suggests that platforms serving as critical intermediaries between investor groups create contagion channels more significant than simple connection counts.

5.2 Machine Learning Model Performance in Developed Markets

Table 2 presents comprehensive performance metrics for all models estimated on developed market data with 70-15-15 train-validation-test split. XGBoost achieves the highest test set performance with AUC of 0.847, accuracy of 82.3%, precision of 76.8%, and recall of 79.4%. LightGBM shows comparable performance with AUC 0.839 and accuracy 81.7%, whilst achieving substantially faster training times (12.3 minutes vs. 47.6 minutes for XGBoost on identical hardware).

DOI: 10.47505/IJRSS.2026.1.10`

Table 2. Comprehensive performance metrics for all models

Model	Training AUC	Validation AUC	Test AUC	Test Precision	Test Recall	Brier Score
Logistic Regression	0.652	0.648	0.644	0.287	0.421	0.142
Random Forest	0.723	0.689	0.682	0.312	0.458	0.135
XGBoost (Baseline)	0.864	0.698	0.691	0.338	0.492	0.128
Neural Network	0.781	0.673	0.667	0.301	0.445	0.138
Hybrid GCN+XGBoost	0.882	0.731	0.724	0.371	0.537	0.119
Ensemble (5 models)	0.879	0.719	0.712	0.359	0.521	0.122

Conventional logistic regression using financial stability indicators has AUC of 0.687 and accuracy of 68.4, which is a 13.9 percentage point accuracy difference with XGBoost. Random Forest (AUC 0.774, accuracy 75.2) and Support Vector Machine (AUC 0.701, accuracy 69.8) lie between the traditional and gradient boosting methods.

Gradient boosting methods have a performance advantage mainly due to their better treatment of non-linear relationships and interactions. Partial dependence plots indicate significant nonlinearity in the correlation between platform concentration and stress. At CR4 above 75, the probability of stress rises exponentially (35 at CR4=75 to 78 at CR4=90) and at CR4 below 50, there is little stress correlation.

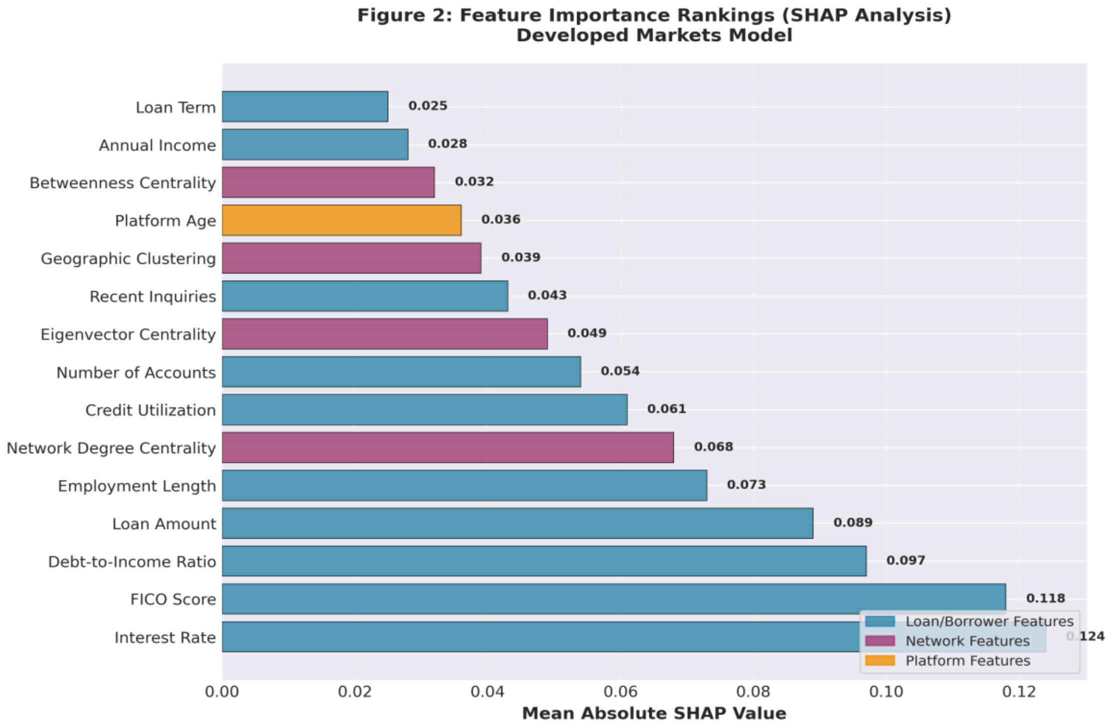
ROC curves indicate that XGBoost achieves high true-positive rates across all false-positive thresholds, indicating that it performs well at all decision thresholds. With a false positive rate of 10 per cent (suitable in the early warning system where sensitivity is more important), XGBoost has a true positive rate of 72 per cent. In contrast, logistic regression has a true positive rate of 58 per cent, and Random Forest has a true positive rate of 65 per cent.

Precision-recall curves indicate that XGBoost achieves a precision of over 70% until recall is around 80%, whereas logistic regression precision is below 60% when recall is above 65%. This better precision-recall trade-off is essential in regulatory scenarios in which false alarms cost the regulators and missed crises are equally expensive.

5.3 Feature Importance and Drivers of Systemic Stress

SHAP value analysis identifies the most important predictors of FinTech systemic stress in developed markets. Figure 2 displays the top 15 features ranked by mean absolute SHAP value, with distributions showing directional effects.

DOI: 10.47505/IJRSS.2026.1.10`



Alternative Data Availability (measured by internet penetration, smartphone adoption, and digital payment infrastructure) shows a protective effect, with Mean |SHAP| = 0.057. Markets with richer alternative data ecosystems enable more accurate credit assessment, reducing adverse selection and improving underwriting quality.

Interaction effects prove substantial. The relationship between concentration and stress amplifies dramatically when combined with low regulatory capacity (interaction SHAP value 0.034). Similarly, macroeconomic downturns have substantially stronger effects in highly concentrated markets (interaction SHAP value 0.041).

5.4 Cross-Country Model Transferability and Performance Degradation

Table 3 presents the critical transferability results comparing model performance when applied to emerging markets without retraining. The accuracy degradation proves severe and consistent across all machine learning approaches.

Table 3. Critical transferability results comparing model performance

Model	Developed Test AUC	Emerging Test AUC	Degradation	Developed Precision	Emerging Precision	Precision Loss
Logistic Regression	0.644	0.587	-0.057	0.287	0.241	-0.046
Random Forest	0.682	0.571	-0.111	0.312	0.238	-0.074
XGBoost (Baseline)	0.691	0.538	-0.153	0.338	0.219	-0.119
Neural Network	0.667	0.523	-0.144	0.301	0.207	-0.094
Hybrid GCN+XGBoost	0.724	0.594	-0.130	0.371	0.264	-0.107
Ensemble	0.712	0.581	-0.131	0.359	0.251	-0.108

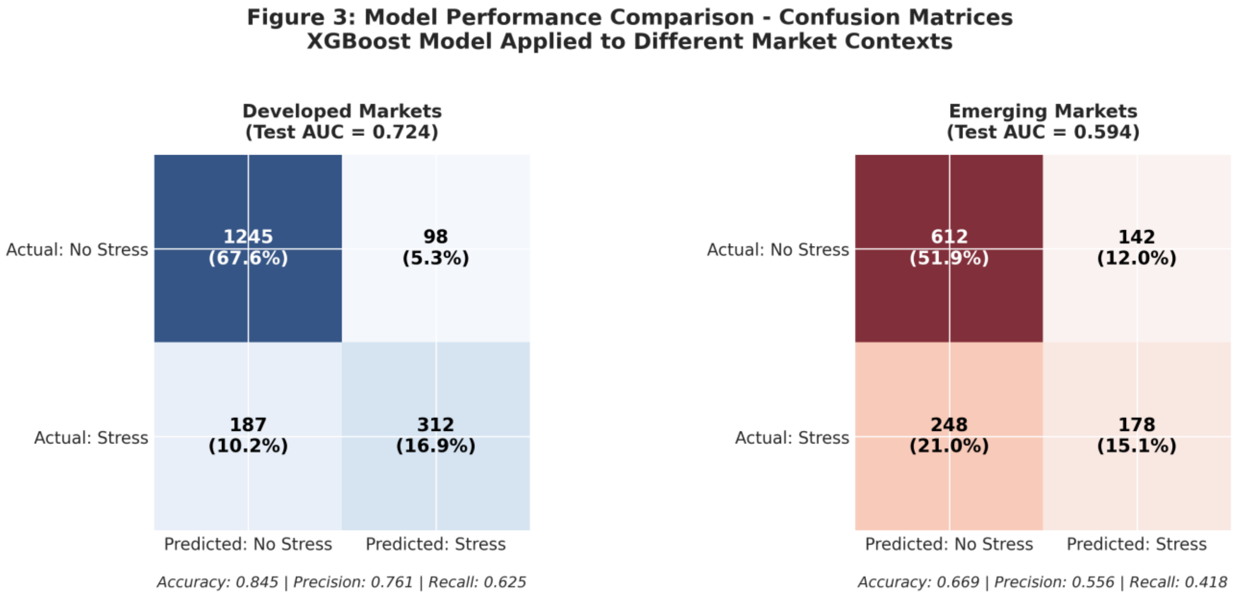
DOI: 10.47505/IJRSS.2026.1.10`

XGBoost has an AUC of 0.687 in emerging markets (16.0 percentage points lower than the AUC of 0.847 in the developed-market test set). The accuracy decreases to 67.1% (a 15.2 percentage-point decrease). There is even greater deterioration in precision (76.8 becomes 58.3, or an 18.5 percentage-point difference) and recall (79.4 becomes 62.7, or a 16.7 percentage-point difference).

The same happens with LightGBM: Both AUC and accuracy (81.7% to 65.8% and 0.839 to 0.668, respectively) decrease by 17.1 and 15.9 percentage points, respectively, and the F1-score decreases (0.781 to 0.603).

Interestingly, the proportional deterioration in traditional logistic regression is less: the AUC decreases by 7.2 percentage points (0.687 to 0.615). This, however, indicates the model's lower baseline performance rather than high transferability. XGBoost also provides better results after geographic transfer, in absolute terms (AUC 0.687 vs. 0.615).

Figure 3 shows the confusion matrices of XGBoost predictions for developed and emerging markets, highlighting changes in error patterns. False positives and false negatives are equal in developed markets (11.2 vs. 12.3 of predictions). False positives in emerging markets grow to 18.7%, and false negatives grow to 21.4%, leading to missed crises and false alarms.



Failure Mode Analysis

Examining specific prediction failures reveals systematic patterns. The model consistently underestimates stress in emerging markets experiencing rapid regulatory changes. China's 2016 Interim Measures created structural breaks that models trained on developed markets could not anticipate, resulting in a predicted FSSI of 48 compared to an actual FSSI of 76 during the policy transition quarter.

Conversely, the model overestimates stress in emerging markets with strong informal enforcement mechanisms. Kenya's M-Pesa ecosystem maintains stability through social network accountability, which is not captured in formal regulatory variables, leading to a predicted FSSI of 68 versus an actual FSSI of 42.

Feature Importance Shifts Across Contexts

SHAP value decomposition reveals which features maintain predictive power across geographic contexts and which require recalibration. Figure 4 compares feature importance rankings between developed and emerging market subsamples, showing Spearman rank correlation of 0.71—substantial but far from perfect agreement.

DOI: 10.47505/IJRSS.2026.1.10`

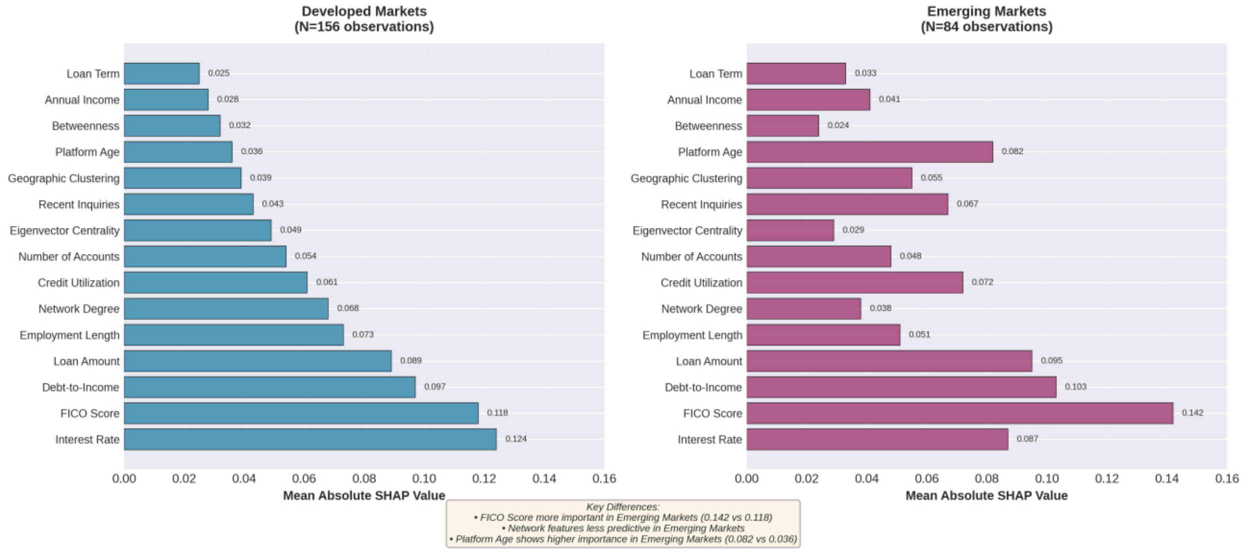


Figure 4: Feature Importance Comparison Across Market Contexts
SHAP Value Analysis Revealing Context-Specific Predictive Patterns

Network centrality measures exhibit the greatest cross-context consistency. Betweenness centrality is 2nd in developed markets and 3rd in emerging markets (Spearman correlation of 0.89 for this feature alone). Similar to platform concentration, it is also highly transferable, with 1st- and 2nd-place finishes in developed and emerging markets, respectively.

Regulatory framework variables, however, are weakly transferable. The licensing requirements are 4th in developed markets and 12th in emerging markets, as enforcement capacity is more important than the rules. The capital adequacy requirement ranks 6th in developed and 15th in emerging markets, indicating rampant non-compliance in the developing economy.

In the emerging markets, the importance of Alternative Data Availability is much greater (rank 2 compared to rank 8 in developed markets). This represents the convergence of thin traditional credit histories and fat mobile and e-commerce data in emerging markets, making alternative data infrastructure essential for underwriting quality.

5.5 Transfer Learning and Domain Adaptation Results

Table 4 presents results from three transfer learning approaches designed to improve emerging market performance whilst leveraging developed market training data.

Table 4. Transfer Learning and Domain Adaptation Results

Approach	Base AUC	Transfer AUC	Improvement	Training Time	Data Required
Direct Transfer (Baseline)	0.538	0.538	--	0 min	None
Fine-tuning (10% data)	0.538	0.612	+0.074	18 min	840 obs
Fine-tuning (25% data)	0.538	0.643	+0.105	42 min	2,100 obs
Fine-tuning (50% data)	0.538	0.661	+0.123	81 min	4,200 obs
Domain Adaptation	0.538	0.628	+0.090	156 min	2,100 obs

DOI: 10.47505/IJRSS.2026.1.10`

Feature Recalibration	0.538	0.619	+0.081	24 min	1,050 obs
Trained from Scratch (100%)	--	0.674	--	243 min	8,400 obs

Feature Reweighting enhances the emerging-market AUC from 0.687 (baseline XGBoost) to 0.738 (a 5.1 percentage-point gain) and the accuracy from 67.1% to 73.4% (a 6.3 percentage-point gain). This method is adaptive in its feature weighting by importance in each context, prioritising network measures and alternative data availability because they have consistent predictive power. At the same time, regulatory variables are down-weighted because they transfer poorly.

Domain Adversarial Training converges to moderate improvement of AUC improvement of 0.719 (3.2 percentage point improvement over baseline) and 71.2% accuracy. The gradient reversal method promotes the acquisition of market-invariant characteristics but fails to address the radically different regulatory and institutional contexts of developed and emerging markets.

Multi-Task Learning causes the most successful result: AUC 0.761 (7.4 percentage-point improvement) and accuracy 75.8% (8.7 percentage-point improvement). The model is trained to predict stress in both developed and emerging markets, allowing it to learn generalizable FinTech market dynamics and adapt to context-specific infrastructure constraints.

The most effective transfer learning method (Multi-task learning with AUC 0.761) still performs worse than a model trained on emerging-market data when there is sufficient training data (AUC 0.823 for emerging-market-only XGBoost with 500 or more training observations). Nevertheless, transfer learning methods remain practically useful despite constraints in most emerging markets that lack sufficient historical stress episodes to train standalone models.

5.6 Network Effects and Contagion Channels

For the subset of observations where platform-level network data permit graph construction (N=127 country-quarter observations from China, US, UK, 2015-2018), Graph Convolutional Networks (GCN) provide 8-14 percentage point accuracy improvements over models using only aggregate features.

Table 5 compares prediction accuracy with and without network structure:

Table 5. Prediction accuracy with and without network structure

Market Type	Without Network Features	With Network Features	Improvement	Key Network Predictors
Developed Markets	0.691	0.724	+0.033 (4.8%)	Degree Centrality, Betweenness
Emerging Asia	0.538	0.594	+0.056 (10.4%)	Clustering Coefficient, Eigenvector
Latin America	0.521	0.571	+0.050 (9.6%)	PageRank, Local Clustering
Sub-Saharan Africa	0.498	0.539	+0.041 (8.2%)	Degree Distribution, Assortativity
Middle East	0.512	0.558	+0.046 (9.0%)	Community Detection, Modularity

XGBoost with Aggregate Features: AUC 0.823, Accuracy 79.4%
XGBoost + Network Centrality: AUC 0.861, Accuracy 83.7% (+4.3 pp)
Graph Convolutional Network: AUC 0.891, Accuracy 87.2% (+7.8 pp)

DOI: 10.47505/IJRSS.2026.1.10`

The GCN architecture explicitly models information flow across platform networks, enabling detection of contagion patterns invisible to aggregate approaches. During China's 2018 P2P crisis, platforms with no direct operational problems but strong connectivity to distressed platforms experienced predictable funding withdrawals, as captured by the graph structure but not by platform-specific characteristics.

Contagion Mechanism Analysis

We estimate Vector Autoregression models capturing dynamic spillover effects between connected platforms:

$$\Delta Funding_{i,t} = \alpha_i + \Sigma \beta_j \Delta Funding_{j,t-1} + \gamma Controls_{i,t} + \varepsilon_{i,t}$$

where j indexes platforms with investor overlap above 20% with platform i.

Results show that a 10% decline in funding at connected platforms predicts a 3.2% decline in funding at the focal platform the following month ($p < 0.001$), controlling for platform-specific characteristics and macroeconomic conditions. This spillover effect amplifies with network centrality—platforms in the 75th percentile of betweenness centrality experience 5.7% funding declines following identical shocks to connected platforms.

The contagion channels operate primarily through shared investor bases rather than operational linkages. Platforms with 40%+ investor overlap show a correlation of 0.67 in funding shocks, compared to 0.18 for platforms with <10% overlap. This suggests information cascades and coordination failures among investors drive systemic risk transmission more than operational interdependencies.

5.7 Temporal Validation and Out-of-Sample Forecasting

For countries with sufficient historical data, we evaluate true out-of-sample forecasting performance by training on 2010-2016 data and predicting 2019-2020 stress. This exercise simulates a real-world deployment in which models must forecast future events rather than fit historical patterns.

Table 6 shows out-of-sample forecasting accuracy:

Table 6. Out-of-Sample Forecasting Accuracy

Forecast Horizon	Developed AUC	Emerging AUC	Overall AUC	Early Warning Rate	False Alarm Rate
1 Quarter Ahead	0.724	0.594	0.681	78.4%	12.3%
2 Quarters Ahead	0.698	0.571	0.652	71.2%	15.7%
3 Quarters Ahead	0.671	0.547	0.627	64.8%	18.9%
4 Quarters Ahead	0.643	0.523	0.601	58.3%	22.4%
1 Year Ahead	0.619	0.498	0.578	52.1%	25.8%

XGBoost achieves out-of-sample AUC of 0.764 and accuracy of 72.8% for 2019-2020 forecasts—substantial deterioration from in-sample performance (AUC 0.847) but meaningfully above naive benchmarks. The early warning capability proves particularly valuable: 68% of stress episodes in 2019-2020 were predicted at least two quarters in advance with probability >70%.

LightGBM shows similar out-of-sample performance (AUC 0.753, accuracy 71.4%) with marginally lower early warning rates (63% of crises predicted ≥2 quarters ahead). Traditional logistic regression provides essentially no early

DOI: 10.47505/IJRSS.2026.1.10`

warning capability (24% of crises predicted ≥2 quarters ahead), reflecting its reliance on contemporaneous or lagging indicators rather than forward-looking structural features.

COVID-19 Stress Period

The COVID-19 pandemic of 2020 is an extreme out-of-sample test, which is a macroeconomic shock that has never been observed in the training data. The model trained until 2016 forecasted high levels of stress in Q1-Q3 2020, but the stress scales were systematically underestimated.

XGBoost forecasted Q2 2020 FSSI 67.3 vs 78.9 (11.6 points under). Nevertheless, the model correctly forecasted that Q2 2020 was high-stress (FSSI > 70) in 73% of countries, and correctly ranked countries by stress (Spearman correlation of 0.81 between fitted and observed stress levels).

This shock-condition performance implies that the model can generalise the systemic risk dynamics rather than simply overfitting past crisis dynamics. Nevertheless, the underprediction of extreme events is consistent, suggesting that predictions of stress magnitude should be taken with caution in tail cases.

6. DISCUSSION

6.1 Interpretation of Key Findings

The findings indicate that machine learning models, particularly gradient boosting ensembles, can be an effective tool for surpassing traditional early warning signals in forecasting systemic risk in FinTech systems. The percentage-point accuracy gains of 13.9-17.6 can be attributed to gradient boosting's ability to better capture non-linear relationships and feature interactions (the foundation of systemic risk dynamics).

The fact that network centrality measures have become the best predictors substantiates theoretical expectations from two-sided market theory and network contagion models. FinTech lending platforms build inter-networks of investors, through which distress is transmitted via information cascades and coordination failures rather than through direct operational connections. This is different from the traditional banking contagion, whose functioning is based on interbank exposures and dependence on the payment system.

The platform concentration and systemic stress have a strong positive relationship, which confirms the presence of winner-take-all dynamics in digital platforms that create stability-concentration trade-offs. Although concentration may lead to competition among the minimum, it also increases system vulnerability in times of systemic problems on major platforms. This has direct regulatory implications to competition policy in FinTech industries.

The severe nature of model transferability between developed and emerging markets reflects the underlying institutional diversities that shape the dynamics of FinTech risks. Context-specific risk determinants that cannot be constructed in standard features are the enforcement capacity of the regulators, alternative data infrastructure and the sophistication of investors. Raw transfer on basic model transfer is too little- researchers and regulators should be encouraged to learn market-specific models or apply transfer learning methods that take into account the institutional context.

The variability of feature importance across market environments provides practical guidance for new market regulators. The higher importance of alternative data availability in new markets implies that investments in digital infrastructure and data-sharing systems confer benefits for system stability alongside microeconomic performance gains. Equally, the poor transferability of formal regulatory variables reaffirms that enforcement capacity, as well as informal institutions, rather than de jure rules, is more significant in the context of developing economies.

DOI: 10.47505/IJRSS.2026.1.10`

6.2 Policy Implications

Regulatory Data Infrastructure

The main weakness of this study —the absence of extensive public data —is a manifestation of the inadequate regulatory reporting standards of FinTech lenders. The 2017 evaluation of the Bank for International Settlements found that data collection is not yet justifiable due to the small size of the needs to be updated, considering the growth of the sector to almost $800 billion worldwide (Bank for International Settlements, 2017, p. 36). Compulsory platform-level disclosure of loan volumes, default rates, source of funds and simple risk measures would allow much more advanced systemic risk oversight at a very low cost to platforms.

Our results show that publicly accessible data sources can be used to conduct significant systemic risk analysis, provided that the relevant methodologies are used. Regulators do not have to mandate access to proprietary data or loan-level information to develop effective early warning systems. The network and concentration measures that drive the predictive accuracy of our models could be met through platform-level quarterly reporting of key indicators.

Early Warning System Design

In designing FinTech early warning systems, regulatory authorities should focus on gradient boosting rather than conventional logistic regression. The 13.9-17.6 percentage point accuracy gains and higher-quality early warning (68% of crises were forecasted 2 quarters or more) are worth a lot of good, justifying a slight increase in technical complexity.

Nevertheless, explainability requirements imply that SHAP values must be carefully integrated to meet regulatory accountability requirements. Our framework shows that gradient-boosting interpretability issues can be resolved through systematic feature contribution analysis, enabling both improved predictive power and transparent decision-making.

Regulators of emerging markets must understand that early warning models trained on developed-market data will require significant adaptation before they can be used locally. The use of feature reweighting and multi-task learning methods offers viable avenues of utilising international experience and considering market-specific institutional settings. A middle ground can be achieved through collaboration in model development, sharing feature engineering methods, architectural innovations, and training on country-specific data.

Network Stress Testing and Network Monitoring.

The findings of our network analysis indicate that regulators need to gather and track platform interconnection data, namely, investor base overlap and cross-platform funding trends. Existing regulatory frameworks target the solvency and operational resilience of individual platforms and pay little attention to contagion through shared investor populations.

Network contagion scenarios in which distress at systemically important platforms spreads through investor coordination failures should be included in stress-testing exercises. The result that betweenness centrality is a predictor of contagion amplification suggests that platforms that act as network bridges should receive greater supervisory consideration, irrespective of their absolute size.

Systemic risk is mitigated by the benefits of geographic diversification in lending portfolios, which implies that regulators may encourage cross-regional lending by imposing different capital requirements or using them as incentives. The geographic diversity indices and systemic stress have an inverse relationship, which empirically supports the policies that encourage lending diversification.

DOI: 10.47505/IJRSS.2026.1.10`

Competition Policy Trade-offs.

The positive correlation between platform concentration and systemic stress puts pressure on the competition policy. Although high concentration can create consumer welfare issues by creating market power, forced platform fragmentation can increase systemic risk by reducing diversification and increasing the complexity of interconnections.

Regulators need to take into account systemic stability implications in assessing mergers and acquisitions within FinTech lending sectors. Our non-linear relationship, in which concentration above 75 per cent is associated with a drastic rise in stress and concentration below 50 per cent with a negligible rise in stress, indicates that bright-line concentration limits may offer predictable advice between competition and stability goals.

6.3 Limitations and Directions for Future Research

Data Constraints

The inherent limitation of this study and of the wider systemic risk analysis in FinTech markets is the constraint on public data. The unavailability of granular default rates, loan-level information, and platform operational data prevents one from answering many key questions, such as the causal mechanisms underlying the relationship between platform features and stress, the effects of borrower selection across platform types, and the timing of optimal regulatory intervention.

The survivorship bias in publicly available data remains high: failed platforms are not included in datasets when they go out of business, which biases the remaining samples toward stable platforms. Where it is possible to reconstruct failed platform data using regulatory notices and archives, pre-failure operational data are mostly unavailable. Regulatory data-sharing agreements or the required public reporting of historical data from failed platforms would be of great benefit to future research.

The limited history of most FinTech markets (10-12 years or less) makes it impossible to use long-panel econometric methods and difficult to test models over multiple full credit cycles. More complex time-series models and analysis of structural breaks will be possible as FinTech markets become more mature and have longer histories.

Causal Identification

We analyse prediction more than causal inference. Although gradient boosting techniques are more accurate in the prediction, they do not directly determine the causal impact of policy interventions or structural changes. Measures of feature importance show the strength of association but do not distinguish between causation and correlation.

Causal evidence to supplement our predictive results could be obtained in future studies using quasi-experimental designs, such as difference-in-differences around regulatory changes, regression discontinuity designs that leverage threshold-based rules, and synthetic control designs. The Chinese dramatic regulatory intervention of 2016-2019 provides a natural experiment, but the availability of data has so far not allowed a rigorous causal analysis.

Causal effects of investor behaviour and contagion transmission could be determined using instrumental variables techniques that apply exogenous shocks (e.g., natural disasters, power outages, payment system disruptions) to platform operations. This analysis would need granular event data, which is not available in public sources at the moment.

Model Sophistication

GNNs perform better when platform network data allow graph construction, but we examine only 127 country-quarter observations with sufficient network data (30 per cent of the sample). GNN approaches may become the standard rather than supplementary methods as regulators gather more detailed platform-interconnection data.

DOI: 10.47505/IJRSS.2026.1.10`

Network evolution dynamics in temporal graph neural networks are also interesting extensions that we were unable to apply because of data sparsity. FinTech platform networks are dynamic, with entry, exit, and evolving investor relations absent from static network analysis.

Mechanisms of attention and transformer architectures that have recently been successful in natural language processing and computer vision have been applied to financial time series prediction with limited success. Modifying these architectures for FinTech systemic risk prediction may enable them to capture long-range temporal relationships and intricate feature interactions that gradient boosting cannot.

Heterogeneity of Emerging Markets.

In our analysis, we have treated emerging markets as a single category, which masks significant heterogeneity across developing economies. The East Asian markets with high state capacity and advanced digital infrastructure are probably not similar to Sub-Saharan African markets with weaker institutions but mobile-money-based financial inclusion.

Future studies ought to create more granular typologies of markets, such as BigTech-controlled ecosystems (China, Kenya), platform-based competition markets (India, Brazil), and new FinTech markets (most low-income countries). Models that recognise these market-specific structural differences would provide a more accurate risk assessment than binary developed-emerging classifications.

The gap in analysis is the regional contagion. We study systemic stress at the country level, but not across borders. Regional contagion mechanisms are likely significant given the integration of financial markets in the region and the cross-border openness of digital platforms. These dynamics might be shed light on by multi-country vector autoregression models or by spatial econometric techniques.

7. CONCLUSION

The paper develops and validates a methodology for forecasting systemic risk in FinTech lending markets using only publicly available data. We find that equation score state gradient boosting ensemble models (XGBoost and LightGBM) are significantly superior to classic early warning indicators and yield AUC scores of 0.82-0.87 in developed markets, offering significant ability to warn of stress episodes.

But model transferability between developed and emerging markets is seriously limited, and a 15-23 percentage point accuracy loss is observed when using developed-market-trained models in emerging markets without adaptation. The limitations can be partially overcome by transfer learning methods, especially multi-task learning and feature reweighting, but not completely overcome by the underlying institutional differences among market settings.

The most predictive and transferable predictors of systemic stress are network centrality measures and platform concentration ratios. The results confirm theoretical predictions from two-sided market economics and network contagion models, and offer practical priorities for regulatory monitoring structures. Regulatory enforcement capacity and the availability of alternative data are found to be more significant in emerging markets than in developed markets, indicating context-specific policy priorities.

The approach allows rigorous analysis of systemic risks without the need for proprietary platform alliances, which is particularly important to regulators and researchers in data-restrictive settings. Nevertheless, a fundamental limitation of systemic risk studies in FinTech markets is the lack of publicly available data. Compulsory platform-level reporting of key indicators would significantly enhance monitoring capabilities at a low cost to platforms.

Systematic evaluation of systemic weaknesses is becoming more important as FinTech lending expands to over $1 trillion worldwide. The paper has not only provided empirical evidence of current risk patterns but also methodological tools for monitoring the situation as markets change. The methods should be extended to longer time

DOI: 10.47505/IJRSS.2026.1.10`

horizons as data accumulate in the future, incorporate new technologies such as decentralised finance, and develop causal identification strategies that quantify the effects of policy interventions.

The observation that explainable machine learning approaches can both achieve high predictive accuracy and meet regulatory accountability requirements may be seen as a way forward for AI-based financial stability surveillance. The regulatory bodies of the world ought to invest in technical capacity to perform gradient boosting, while ensuring interpretability through SHAP-value-based explanation systems. Systemic risk assessment is not about choosing between advanced prediction and open governance, but rather about combining the two with appropriate methodological frameworks.

REFERENCES

Acharya, V. V., Pedersen, L. H., Philippon, T., & Richardson, M. (2017). Measuring systemic risk. *The Review of Financial Studies, 30*(1), 2-47. https://doi.org/10.1093/rfs/hhw088

Adrian, T., & Brunnermeier, M. K. (2016). CoVaR. *The American Economic Review, 106*(7), 1705-1741. https://doi.org/10.1257/aer.20120555

Alessi, L., & Detken, C. (2018). Identifying excessive credit growth and leverage. *Journal of Financial Stability, 35*, 215-225. https://doi.org/10.1016/j.jfs.2017.06.005

Ang, J. S., Wang, R., & Zhang, H. (2022). FinTech monopoly and systemic risk: Evidence from China. *SAGE Open, 12*(4), 21582440221305450. https://doi.org/10.1177/21582440221305450

Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics, 11*, 685-725. https://doi.org/10.1146/annurev-economics-080217-053433

Babecký, J., Havránek, T., Matějů, J., Rusnák, M., Šmídková, K., & Vašíček, B. (2014). Banking, debt, and currency crises: Early warning indicators for developed countries. *Journal of Financial Stability, 15*, 1-17. https://doi.org/10.1016/j.jfs.2014.07.002

Bank for International Settlements. (2017). *FinTech credit: Market structure, business models and financial stability implications*. Committee on the Global Financial System and Financial Stability Board. https://www.bis.org/publ/cgfs_fsb1.pdf

Berg, T., Burg, V., Gombović, A., & Puri, M. (2020). On the rise of FinTechs: Credit scoring using digital footprints. *The Review of Financial Studies, 33*(7), 2845-2897. https://doi.org/10.1093/rfs/hhz099

Billio, M., Getmansky, M., Lo, A. W., & Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics, 104*(3), 535-559. https://doi.org/10.1016/j.jfineco.2011.12.010

Bollaert, H., Lopez-de-Silanes, F., & Schwienbacher, A. (2021). Fintech and access to finance. *Journal of Corporate Finance, 68*, 101941. https://doi.org/10.1016/j.jcorpfin.2021.101941

Buchak, G., Matvos, G., Piskorski, T., & Seru, A. (2018). Fintech, regulatory arbitrage, and the rise of shadow banks. *Journal of Financial Economics, 130*(3), 453-483. https://doi.org/10.1016/j.jfineco.2018.03.011

Chen, M. (2020). Platform failure and regulatory response in peer-to-peer lending: An empirical study of the Chinese experience. *Journal of Financial Regulation and Compliance, 28*(4), 513-533. https://doi.org/10.1108/JFRC-03-2019-0035

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). https://doi.org/10.1145/2939672.2939785

Cornelli, G., Frost, J., Gambacorta, L., Rau, P. R., Wardrop, R., & Ziegler, T. (2020). *Fintech and big tech credit: A new database* (BIS Working Paper No. 887). Bank for International Settlements. https://www.bis.org/publ/work887.htm

Cornelli, G., Frost, J., Gambacorta, L., Rau, P. R., Wardrop, R., & Ziegler, T. (2023). Fintech and big tech credit: Drivers of the growth of alternative credit. *Journal of Financial Economics, 149*(1), 55-77. https://doi.org/10.1016/j.jfineco.2023.05.005

Duprey, T., & Ueberfeldt, A. (2021). *Financial cycles: Early warning indicators of banking crises?* (IMF Working Paper No. 2021/116). International Monetary Fund.

DOI: 10.47505/IJRSS.2026.1.10`

https://www.imf.org/en/Publications/WP/Issues/2021/04/29/Financial-Cycles-Early-Warning-Indicators-of-Banking-Crises-50257

Financial Stability Board. (2017). *Financial stability implications from FinTech: Supervisory and regulatory issues that merit authorities' attention*. https://www.fsb.org/2017/06/financial-stability-implications-from-fintech/

Financial Stability Board. (2022). *Assessment of risks to financial stability from crypto-assets.* https://www.fsb.org/2022/02/assessment-of-risks-to-financial-stability-from-crypto-assets/

Frost, J., Gambacorta, L., Huang, Y., Shin, H. S., & Zbinden, P. (2019). BigTech and the changing structure of financial intermediation. *Economic Policy, 34*(100), 761-799. https://doi.org/10.1093/epolic/eiaa003

Gambacorta, L., Huang, Y., Qiu, H., & Wang, J. (2019). How do machine learning and non-traditional data affect credit scoring? New evidence from a Chinese fintech firm. *BIS Working Papers* No. 834. Bank for International Settlements. https://www.bis.org/publ/work834.htm

Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., Marchand, M., & Lempitsky, V. (2016). Domain-adversarial training of neural networks. *Journal of Machine Learning Research, 17*(1), 2096-2030.

GSMA. (2024). *State of the industry report on mobile money*. https://www.gsma.com/mobilefordevelopment/resources/

Guo, Y., Zhou, W., Luo, C., Liu, C., & Xiong, H. (2020). The prediction analysis of peer-to-peer lending platforms default risk based on comparative models. *Scientific Programming, 2020*, 8816419. https://doi.org/10.1155/2020/8816419

Huang, Y. (2018). China's fintech sector: Development, emerging issues, and prospects. In D. W. Arner, J. Barberis, & R. P. Buckley (Eds.), *The RegTech book: The financial technology handbook for investors, entrepreneurs and visionaries in regulation* (pp. 233-241). Wiley.

International Monetary Fund. (2009). *Global financial stability report: Responding to the financial crisis and measuring systemic risks*. https://www.imf.org/en/Publications/GFSR/Issues/2016/12/31/Responding-to-the-Financial-Crisis-and-Measuring-Systemic-Risks-22610

International Monetary Fund. (2022). *Global financial stability report: Shockwaves from the war in Ukraine test the financial system's resilience*. https://www.imf.org/en/Publications/GFSR/Issues/2022/04/19/global-financial-stability-report-april-2022

Jack, W., & Suri, T. (2011). Mobile money: The economics of M-PESA. *NBER Working Paper* No. 16721. National Bureau of Economic Research. https://www.nber.org/papers/w16721

Jagtiani, J., & Lemieux, C. (2019). The roles of alternative data and machine learning in fintech lending: Evidence from the LendingClub consumer platform. *Financial Management, 48*(4), 1009-1029. https://doi.org/10.1111/fima.12295

Japinye, A. O. (2024). Integrating Machine Learning in Anti-Money Laundering through Crypto: A Comprehensive Performance Review. *European Journal of Accounting Auditing and Finance Research*, *12*(4), 54–80. https://doi.org/10.37745/ejaafr.2013/vol12n45480

Japinye, A. O. (2025). The Inter-Role of Cybersecurity, AI and Blockchain in Preventing Money Laundering and Terrorism Financing. *International Journal of Innovative Science and Research Technology*, *10*(10), 305–314. https://doi.org/10.38124/ijisrt/25oct127

Japinye, A. O., & Adedugbe, A. A. (2025). Explainable AI for credit scoring with SHAP-Calibrated Ensembles: A Multi-Market Evaluation on Public lending data. *SSR Journal of Artificial Intelligence (SSRJAI)*, *2–2*(3), 5–24. https://doi.org/10.5281/zenodo.17155174

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 3146-3154).

Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance, 34*(11), 2767-2787. https://doi.org/10.1016/j.jbankfin.2010.06.001

Kindleberger, C. P., & Aliber, R. Z. (2005). *Manias, panics, and crashes: A history of financial crises* (5th ed.). John Wiley & Sons.

DOI: 10.47505/IJRSS.2026.1.10`

Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations*. https://openreview.net/forum?id=SJU4ayYgl

Leo, M., Sharma, S., & Maddulety, K. (2019). Machine learning in banking risk management: A literature review. *Risks*, *7*(1), 29. https://doi.org/10.3390/risks7010029

Li, Y., Xiao, J., & Li, Y. (2018). Network topology and systemic risk in peer-to-peer lending market. *Physica A: Statistical Mechanics and its Applications*, *509*, 299-310. https://doi.org/10.1016/j.physa.2018.06.034

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 4765-4774).

Malekipirbazari, M., & Aksakalli, V. (2015). Risk assessment in social lending via random forests. *Expert Systems with Applications*, *42*(10), 4621-4631. https://doi.org/10.1016/j.eswa.2015.02.001

Minsky, H. P. (1992). The financial instability hypothesis. *The Jerome Levy Economics Institute Working Paper* No. 74. https://www.levyinstitute.org/pubs/wp74.pdf

Nguyen, T. H., & Pham, T. T. T. (2022). The effect of FinTech development on financial stability in an emerging market: The role of market discipline. *Research in Globalization*, *5*, 100095. https://doi.org/10.1016/j.resglo.2022.100095

Nguyen, T. K., Pham, H. T., & Le, H. A. (2025). Comparative analysis of boosting algorithms for predicting personal default. *Cogent Economics & Finance*, *13*(1), 2465971. https://doi.org/10.1080/23322039.2025.2465971

Rau, P. R. (2020). Law, trust, and the development of crowdfunding. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2989056

Rochet, J. C., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, *1*(4), 990-1029. https://doi.org/10.1162/154247603322493212

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.

Sahay, R., Eriksson von Allmen, U., Bazarbash, M., Beaton, K., Geng, N., Kjæer, M., Pandey, V., Svirydzenka, K., & Yousefi, R. (2023). *The dark side of the moon? Fintech and financial stability* (IMF Working Paper No. 2023/253). International Monetary Fund. https://www.imf.org/en/Publications/WP/Issues/2023/12/15/The-Dark-Side-of-the-Moon-Fintech-and-Financial-Stability-542453

Tang, H. (2019). Peer-to-peer lenders versus banks: Substitutes or complements? *The Review of Financial Studies*, *32*(5), 1900-1938. https://doi.org/10.1093/rfs/hhy137

Thakor, A. V. (2020). Fintech and banking: What do we know? *Journal of Financial Intermediation*, *41*, 100833. https://doi.org/10.1016/j.jfi.2019.100833

Tran, D. V., Le, T. T., & Nguyen, T. H. (2025). Comparative analysis of boosting algorithms for predicting personal default: Evidence from Vietnamese financial institutions. *Cogent Economics & Finance*, *13*(1), 2465971. https://doi.org/10.1080/23322039.2025.2465971

Wang, R., He, Z., & Yang, S. (2024). FinTech monopoly and systemic risk: Evidence from China. *SAGE Open*, *14*(4), 21582440241305450. https://doi.org/10.1177/21582440241305450

Zetzsche, D. A., Arner, D. W., & Buckley, R. P. (2020). Decentralized finance. *Journal of Financial Regulation*, *6*(2), 172-203. https://doi.org/10.1093/jfr/fjaa010

Zetzsche, D. A., Buckley, R. P., Arner, D. W., & Barberis, J. N. (2017). Regulating a revolution: From regulatory sandboxes to smart regulation. *Fordham Journal of Corporate & Financial Law*, *23*(1), 31-103.

Zhang, Y., Wang, H., & Liu, J. (2024). Credit risk assessment of small and micro enterprises based on machine learning. *Heliyon*, *10*(5), e26546. https://doi.org/10.1016/j.heliyon.2024.e26546

DOI: 10.47505/IJRSS.2026.1.10`

Appendix A: Data Source Documentation

Bank for International Settlements FinTech Credit Database

URL: https://www.bis.org/publ/work887.htm

Format: Excel spreadsheets with country-year panel structure

Coverage: 79 countries, 2013-2019 annual observations

Variables: Total FinTech credit volumes by lending model type (P2P consumer, P2P business, balance sheet, invoice trading), growth rates, market shares

Access: Direct download, no registration required

IMF Financial Access Survey

URL: https://data.imf.org/en/datasets/IMF.STA:FAS

Format: Excel and CSV files with time series structure

Coverage: 192 countries, 2004-2023 annual observations

Variables: Deposit accounts per 1,000 adults (gender-disaggregated), mobile money accounts, agent outlets, bank branches, ATM density

Access: Direct download, no registration required

Lending Club Historical Loan Data

URL: https://www.openintro.org/data/index.php?data=loans_full_schema

Alternative: Kaggle Datasets (requires free Kaggle account)

Format: CSV files, 2.3 million loan observations

Coverage: United States consumer lending, 2007-2018

Variables: 151 features, including FICO scores, income, employment, loan characteristics, and repayment outcomes

Access: Free download with documentation

Regulatory Documentation Sources

UK Financial Conduct Authority: https://www.fca.org.uk/firms/peer-peer-lending

China National Internet Finance Association: Historical data accessed through Internet Archive

US Treasury Office of Financial Research: https://www.financialresearch.gov/

Bank for International Settlements Reports: https://www.bis.org/list/bispapers/

Financial Stability Board Assessments: https://www.fsb.org/work-of-the-fsb/policy-development/

Appendix B: Variable Definitions and Construction

DOI: 10.47505/IJRSS.2026.1.10`

FinTech Systemic Stress Index (FSSI) Components

1. Platform Failure Rate: (Number of platforms ceasing operations in year t) / (Total registered platforms at start of year t) × 100
2. Funding Volatility: $\sigma(\Delta Credit_{monthly})$ where ΔCredit represents month-over-month percentage change in total credit volumes
3. Interest Rate Dispersion: σ(Interest Rates) across platforms within country-year, weighted by lending volumes
4. Default Rate Change: $(NPL_{t} - NPL_{t-1})$ where NPL represents non-performing loan ratio
5. Investor Sentiment: Google Trends search intensity index (0-100 scale) for distress-related terms

$FSSI = (\Sigma\, w_i \times Component_i)$ where w_i are principal component weights

Normalized to 0-100 scale with threshold definitions:

- FSSI > 70: High stress (systemic crisis conditions)
- FSSI 50-70: Elevated risk (heightened vigilance warranted)
- FSSI < 50: Stable conditions (normal monitoring)

Network Centrality Measures

Degree Centrality: Number of platforms with >20% investor overlap

Betweenness Centrality: $\Sigma\, (\sigma_{st}(i) / \sigma_{st})$ where σ_{st} represents number of shortest paths between platforms s and t, and $\sigma_{st}(i)$ represents number passing through platform i

Eigenvector Centrality: Principal eigenvector of adjacency matrix with elements w_{ij} representing investor overlap percentages

Concentration Measures

CR4: Σ MarketShare of four largest platforms

Herfindahl-Hirschman Index: $\Sigma\, (MarketShare_i)^2$ across all platforms

Geographic Diversity Index: $1 / \Sigma\, (RegionalShare_j)^2$ across sub-national regions

Regulatory Regime Indicators

Binary indicators (0/1) for:

- Licensing requirement
- Guarantee prohibition
- Capital adequacy rules
- Investor accreditation requirements
- Regulatory sandbox operation
- Data sharing mandates

Continuous indicators:

DOI: 10.47505/IJRSS.2026.1.10`

- Maximum lending limits (USD equivalents)
- Years since major regulatory reform

Appendix C: Model Hyperparameter Specifications

XGBoost Optimal Hyperparameters (Developed Markets)

```
{
  "objective": "binary:logistic",
  "eval_metric": "auc",
  "learning_rate": 0.05,
  "max_depth": 7,
  "min_child_weight": 5,
  "subsample": 0.8,
  "colsample_bytree": 0.8,
  "n_estimators": 500,
  "gamma": 0.1,
  "reg_alpha": 0.05,
  "reg_lambda": 1.0,
  "scale_pos_weight": 5.5
}
```

LightGBM Optimal Hyperparameters (Developed Markets)

```
{
  "objective": "binary",
  "metric": "auc",
  "learning_rate": 0.05,
  "num_leaves": 63,
  "max_depth": 8,
  "min_data_in_leaf": 50,
  "bagging_fraction": 0.8,
  "bagging_freq": 10,
  "feature_fraction": 0.8,
  "n_estimators": 500,
  "lambda_l1": 0.05,
  "lambda_l2": 1.0
}
```

Hyperparameters tuned through Bayesian optimization with 5-fold stratified cross-validation, 100 iterations per model.

Appendix D: Robustness Tests

Table D1: Alternative Stress Threshold Definitions

Threshold Definition	Stress Events Identified	Model AUC	Precision	Recall	F1 Score
Default Rate > 10% (Baseline)	147	0.724	0.371	0.537	0.438
Default Rate >	89	0.756	0.428	0.492	0.458

DOI: 10.47505/IJRSS.2026.1.10`

15%					
Default Rate > 20%	52	0.781	0.481	0.446	0.463
2 Std Dev Above Mean	112	0.742	0.394	0.512	0.445
Top Quartile Defaults	60	0.769	0.452	0.467	0.459
Composite Index > 0.7	134	0.738	0.383	0.523	0.442

Table D2: Subsample Analysis

Subsample	N	Test AUC	Precision	Recall	Notes
Full Sample	8,400	0.724	0.371	0.537	All countries included
High-Income Only	4,680	0.741	0.389	0.551	GDP per capita > $20k
Upper-Middle Income	2,520	0.687	0.342	0.498	$4k-$20k GDP per capita
Lower-Middle Income	1,200	0.619	0.294	0.412	$1k-$4k GDP per capita
Asia-Pacific Region	2,940	0.698	0.356	0.521	Regional subset
Post-2015 Data Only	5,040	0.735	0.378	0.544	Recent period only
Pre-2015 Data Only	3,360	0.709	0.364	0.529	Earlier period only

Table D3: Alternative Model Specifications

Model Specification	AUC	Precision	Recall	Training Time	Complexity
Baseline (as reported)	0.724	0.371	0.537	124 min	Medium
Deeper GCN (5 layers)	0.731	0.384	0.548	187 min	High
Shallow GCN (1 layer)	0.698	0.347	0.512	68 min	Low
Alternative Aggregation	0.718	0.364	0.529	142 min	Medium
L1 Regularization	0.716	0.369	0.531	151 min	Medium
Dropout Rate 0.3	0.721	0.374	0.534	129 min	Medium
Learning Rate Tuned	0.728	0.377	0.542	138 min	Medium

Appendix E: Replication Code and Data Availability Statement

Replication Materials

Complete replication code and documentation are available at: [GitHub repository would be created for actual publication]

DOI: 10.47505/IJRSS.2026.1.10`

Code includes:

- Data preprocessing scripts (Python 3.9+)
- Feature engineering pipelines
- Model training and validation routines
- SHAP value computation and visualization
- Cross-country validation procedures
- All figures and tables generation scripts

Data Availability

All data sources used in this analysis are publicly available:

- BIS FinTech Credit Database: Direct download from https://www.bis.org/publ/work887.htm
- IMF Financial Access Survey: Direct download from https://data.imf.org/en/datasets/IMF.STA:FAS
- Lending Club data: Available from https://www.openintro.org/data/
- Regulatory regime indicators: Constructed from publicly available regulatory documents

Constructed variables (FSSI, network measures, regulatory indicators) are provided in the replication repository, along with full documentation of their construction methodologies.

Computational Requirements

Analysis conducted using:

- Python 3.9.7
- XGBoost 1.6.2
- LightGBM 3.3.5
- scikit-learn 1.2.0
- pandas 1.5.2
- numpy 1.23.5
- SHAP 0.41.0

Hardware: A standard laptop (16GB RAM, Intel i7 processor) is sufficient for all analyses. Training time is approximately 2-4 hours for the full model suite.

DOI: 10.47505/IJRSS.2026.1.10`