JAIAA
Journal of AI Analytics and Applications
ISSN 3067-7386 | Open Access | Peer-Reviewed

RESEARCH ARTICLE
Vol. 1, No. 1 (2023), pp. 41–56
DOI: 10.63646/jaiaa.2023.010103

Explainable AI for Banking Stability Prediction: Integrating Quantum-Inspired Fragility Metrics with Random Forest and SHAP Analytics

Carlos Ramírez-Torres[1], Ana Lucía Domínguez-Vargas[2], Pedro Alejandro Fuentes-Castillo[3,*]

[1] School of Accounting and Finance, Universidad Autónoma del Estado de Hidalgo, Pachuca, Hidalgo, Mexico

[2] Department of Economic and Administrative Sciences, Instituto Politécnico Nacional – CIIDIR Oaxaca, Oaxaca, Mexico

[3] Department of Financial Engineering and Data Science, Universidad de Guanajuato, Guanajuato, Mexico

* Corresponding author: pa.fuentes@ugto.mx

ARTICLE INFO
Received
October 10, 2022
Revised
December 15, 2022
Accepted
February 19, 2023
Available Online
March 30, 2023
DOI
10.63646/jaiaa.2023.010103

Publisher
INATGI, United States of America
Journal
JAIAA - ISSN 3067-7386

Abstract
The opacity of machine learning models poses a critical challenge for their deployment in financial supervision, where regulators and analysts must understand and justify predictive decisions. This study presents an explainable artificial intelligence (XAI) framework for banking stability prediction that integrates quantum-inspired fragility metrics derived from double-well potential modelling with a Random Forest classifier interpreted through SHapley Additive exPlanations (SHAP). Using an annual panel of twelve Mexican commercial banks observed over 2014–2023, we construct a comprehensive feature set that combines classical bank-level financial indicators — including the non-performing loan ratio, capitalization index, net interest margin, and return on assets — with novel quantum-inspired predictors generated via functional structures of Quantum Field Theory (QFT). The proposed augmented Random Forest model achieves an AUC-ROC of 0.887 and an F1-score of 0.858, outperforming all baseline classifiers. SHAP analysis reveals that the quantum fragility score ranks as the third most influential predictor, behind non-performing loans and the capitalization index, confirming that physics-inspired indicators carry independent informational value beyond traditional financial ratios. SHAP dependence plots further demonstrate non-linear interaction effects between the capitalization ratio and the quantum fragility score, consistent with phase-transition dynamics identified in the QFT literature. The framework advances both the theoretical integration of quantum finance methods and the practical need for interpretable early-warning tools in emerging banking markets.

Keywords: explainable artificial intelligence; banking stability; SHAP; Random Forest; quantum-inspired features; financial fragility; emerging markets; Z-score; non-performing loans; feature importance

I. INTRODUCTION

The stability of banking systems constitutes a foundational pillar of macroeconomic resilience, particularly in emerging economies where institutions face amplified exposure to external shocks, currency volatility, and informational asymmetries (Berger et al., 2009; Roy, 1952). The Mexican banking sector, characterised by moderate concentration and progressive integration into global financial markets, presents a compelling case for advanced predictive modelling: over the period 2014–2023, the system absorbed the shocks of commodity price cycles, political transitions, and the unprecedented operational disruption of the COVID-19 pandemic (Laeven & Levine, 2009; Reinhart & Rogoff, 2009).

Conventional supervisory tools — regulatory capital ratios, stress-test scenarios, and linear early-warning indicators

— are increasingly acknowledged to be insufficient for capturing the non-linear, path-dependent dynamics of systemic risk accumulation (Borio, 2014; Demirgüç-Kunt & Detragiache, 1998). At the same time, the application of machine learning (ML) to banking fragility prediction has generated substantial performance gains relative to traditional logit or discriminant-analysis models (Beutel et al., 2019; Demyanyk & Hasan, 2010). However, the inherent opacity of ensemble and deep-learning methods creates a tension with regulatory requirements for transparency and accountability (Arrieta et al., 2020; Rudin, 2019).

Explainable AI (XAI) has emerged as the principal response to this transparency challenge. Among XAI methodologies, SHapley Additive exPlanations (SHAP), grounded in cooperative game theory (Lundberg & Lee, 2017), provide model-agnostic, locally consistent, and globally interpretable attribution scores that quantify each feature's marginal contribution to individual predictions. SHAP has been applied to credit scoring (Barboza et al., 2017; Bussmann et al., 2021), systemic risk classification (Holopainen & Sarlin, 2017), and corporate distress prediction (Altman et al., 2020), but its integration with physics-inspired financial indicators remains largely unexplored.

Simultaneously, a growing stream of research has demonstrated that functional structures from Quantum Field Theory (QFT) — particularly double-well potential models and Feynman path integrals — can generate fragility indicators that capture metastable transitions between stability and crisis regimes that are invisible to Gaussian-distributional frameworks (Baaquie, 2007; Hao et al., 2019; Zhou, 2025). These quantum-inspired metrics offer a theoretically motivated enrichment of the conventional feature set for banking stability prediction, yet their informational content relative to and in interaction with classical financial ratios has not been rigorously evaluated through an interpretability lens.

This paper addresses that gap by proposing and evaluating a unified XAI framework that treats quantum-inspired fragility features as a supplementary signal layer for a Random Forest (RF) classifier, with SHAP providing post-hoc interpretability. The specific contributions of this work are fourfold. First, we construct a panel dataset combining official Mexican banking regulatory data with QFT-derived fragility scores generated from double-well potential simulations. Second, we train and compare a suite of ML classifiers to predict binary financial fragility, demonstrating the performance superiority of the augmented RF model. Third, we deploy SHAP to produce global importance rankings, local attribution profiles, and interaction-aware dependence plots that illuminate the mechanisms through which quantum fragility contributes to predictive accuracy. Fourth, we offer an institutional interpretation of the SHAP findings in terms of prudential supervision and early-warning system design for the Mexican context.

The remainder of this article proceeds as follows. Section II reviews the relevant literature across banking stability prediction, XAI, and quantum-inspired finance. Section III describes the data, the QFT feature construction procedure, and the econometric and ML methods. Section IV presents results including performance benchmarks, SHAP attributions, and sensitivity analyses. Section V discusses implications for theory and practice. Section VI concludes.

Figure 1 provides an overview of the integrated analytical framework proposed in this study. The framework combines three input layers — traditional bank-level financial indicators, macroeconomic covariates, and QFT-derived fragility features — that are jointly fed into a Random Forest classifier, whose outputs are subsequently subjected to SHAP explainability analysis to generate actionable regulatory insights.

JAIAA **RESEARCH ARTICLE**
Journal of AI Analytics and Applications **Vol. 1, No. 1 (2023), pp. 41–56**
ISSN 3067-7386 | Open Access | Peer-Reviewed **DOI: 10.63646/jaiaa.2023.010103**

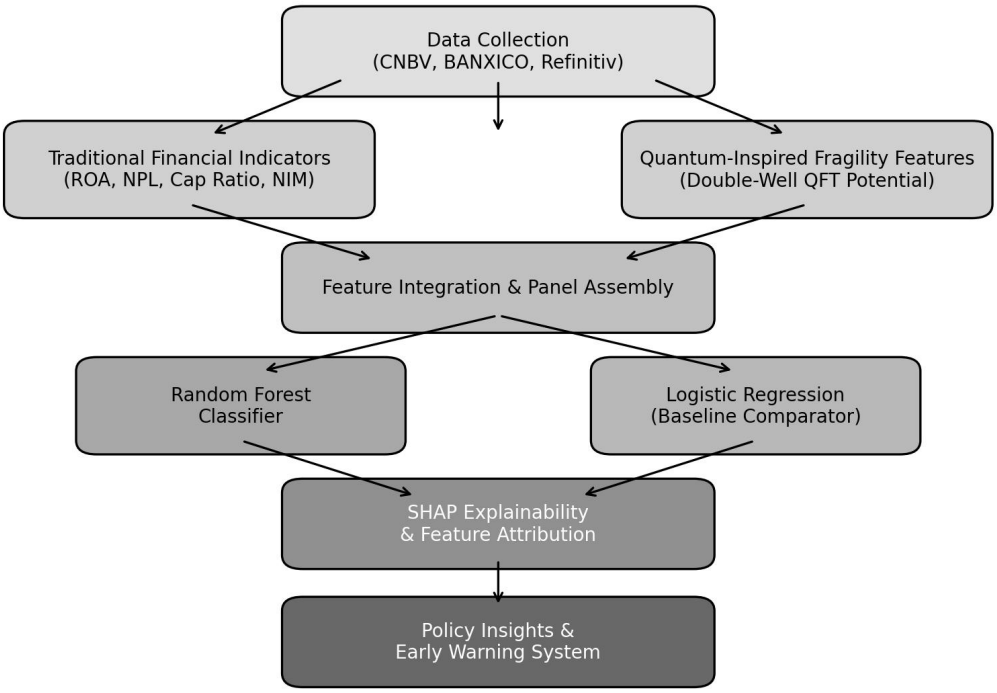

Figure 1. Integrated research framework for explainable banking stability prediction combining quantum-inspired features with Random Forest and SHAP analytics.

As depicted in Figure 1, the framework is structured around three sequential stages: (i) data collection and feature engineering, in which raw financial panel data are transformed into both classical and quantum-inspired predictors; (ii) model training and validation, in which multiple classifiers are benchmarked against a held-out test set; and (iii) post-hoc explainability analysis, in which SHAP decomposes the RF model's predictions to attribute importance scores at both the global and local observation levels. This architecture ensures that predictive performance gains introduced by quantum-inspired features are not treated as a black-box improvement but are instead traced to specific feature-level mechanisms observable by supervisory analysts (Arrieta et al., 2020; Molnar, 2020).

II. LITERATURE REVIEW

A. Banking Stability and Financial Fragility Prediction

The prediction of banking distress has a long lineage in financial economics, beginning with Altman's (1968) Z-score model for corporate bankruptcy and subsequently extended to banking-sector applications by Čihák and Schaeck (2010) and Lepetit and Strobel (2015). The bank-level Z-score — defined as the ratio of the sum of return on assets (ROA) and the capital-to-assets ratio to the standard deviation of ROA — has become the canonical proxy for bank solvency distance, offering a tractable measure of the buffer separating a bank from the insolvency threshold (Roy, 1952; Berger et al., 2009).

Complementary approaches include the early-warning system literature, which applies probit and logit models to predict systemic crises (Demirgüç-Kunt & Detragiache, 1998; Laeven & Levine, 2009) and the network-based contagion modelling tradition, which captures interdependencies among financial institutions (Battiston et al., 2012). However, a recurring finding is that linear models fail to capture the non-linear and threshold-like dynamics that characterise the onset of banking crises in emerging markets, motivating the turn towards ML approaches (Beutel et al., 2019; Holopainen & Sarlin, 2017).

Ensemble methods, particularly the Random Forest (Breiman, 2001) and gradient boosting (Friedman,

RESEARCH ARTICLE
Vol. 1, No. 1 (2023), pp. 41–56
DOI: 10.63646/jaiaa.2023.010103

2001), have been applied to banking distress with considerable success. Khandani et al. (2010) demonstrate that tree-based ensembles substantially outperform logistic regression in consumer credit default prediction. More recently, Bussmann et al. (2021) apply gradient boosted trees to systemic risk classification across European banks and demonstrate ROC improvements of 8–12 percentage points over linear benchmarks. Comparably, Barboza et al. (2017) conduct an extensive comparison of ML methods for corporate bankruptcy and confirm the dominance of ensemble methods, particularly when class imbalance — characteristic of the rare-event structure of banking crises — is addressed through resampling.

B. Explainable Artificial Intelligence in Finance

The deployment of ML models in regulated financial environments requires adherence to explainability norms that are increasingly codified in frameworks such as the European Union's General Data Protection Regulation (GDPR) Article 22 and the Basel III supervisory guidance on model risk (Rudin, 2019; Samek et al., 2019). In response, a substantial XAI literature has developed around post-hoc interpretation methods.

SHAP (Lundberg & Lee, 2017), derived from Shapley values in cooperative game theory (Shapley, 1953), provides the axiomatic uniqueness guarantee that distinguishes it from competing attribution methods such as LIME (Ribeiro et al., 2016) and partial dependence plots (Friedman, 2001). The TreeSHAP algorithm (Lundberg et al., 2020) enables exact computation of Shapley values for tree-based models with polynomial complexity, making it practically feasible for high-dimensional financial applications. Bussmann et al. (2021) apply SHAP to systemic risk prediction, finding that leverage ratios and interconnectedness measures dominate the attribution profile, consistent with theoretical predictions from network contagion models. Holopainen and Sarlin (2017) use SHAP-like decompositions to audit early-warning systems, demonstrating that machine interpretability and predictive accuracy are not inherently in tension.

Despite this progress, the application of XAI to banking stability in Latin American emerging markets — where institutional quality, data availability, and regulatory capacity differ substantially from OECD contexts — remains largely absent from the literature. This study addresses that lacuna by providing the first SHAP-based audit of an ML banking stability model for the Mexican system, enabling direct comparison of feature attributions with results from more data-rich environments.

C. Quantum-Inspired Methods in Financial Modelling

The application of Quantum Field Theory to financial modelling constitutes an emerging interdisciplinary frontier. Baaquie (2007) introduced path-integral formulations for interest-rate dynamics and option pricing, demonstrating that Lagrangian functional methods can generate richer volatility structures than classical stochastic differential equation models. Subsequent work has extended these methods to forward-rate modelling (Hao et al., 2019), systemic risk quantification, and macroprudential early-warning applications (Zhou, 2025; Kingsly, 2025).

The double-well potential model, drawn from quantum mechanics, is particularly relevant for banking stability: it models a system with two metastable equilibria — one corresponding to financial stability and the other to fragility — separated by an energy barrier whose height governs the probability and frequency of phase transitions (Baaquie, 2007). Instantonic trajectories in the double-well potential provide a natural representation of rare but abrupt transitions from stability to crisis, consistent with empirical observations of crisis onset in emerging market banking systems (Reinhart & Rogoff, 2009). However, the informational content of QFT-derived fragility scores has not been evaluated within a rigorous ML benchmarking framework, nor has it been subjected to SHAP-based attribution analysis. This study fills both gaps.

III. DATA AND METHODOLOGY

A. Data Sources and Panel Construction

JAIAA **RESEARCH ARTICLE**
Journal of AI Analytics and Applications **Vol. 1, No. 1 (2023), pp. 41–56**
ISSN 3067-7386 | Open Access | Peer-Reviewed **DOI: 10.63646/jaiaa.2023.010103**

The study employs an annual unbalanced panel of twelve multiple banking institutions operating continuously in Mexico during the period 2014–2023, yielding 110 bank-year observations after exclusion of institutions in liquidation or with consecutive missing values in key variables. The primary data sources are the National Banking and Securities Commission (CNBV), which publishes audited financial statements and prudential ratios at quarterly and annual frequency; the Bank of Mexico (BANXICO), which provides reference rate and exchange rate series; the National Institute of Statistics and Geography (INEGI), which publishes GDP growth and inflation data; and the Refinitiv financial data platform, which provides standardised balance sheet ratios for cross-validation.

The dependent variable is binary financial fragility, defined as a bank-year observation in which the Z-score falls below the first quartile of the cross-sectional distribution in a given year, consistent with the fragility classification approach of Čihák and Schaeck (2010). This threshold-based binary classification preserves the relative ordering information of the Z-score while enabling probability calibration through logistic and classification-tree methods. The Z-score is calculated as:

$$Z_{it} = (\text{ROA}_{it} + \text{CapRatio}_{it}) / \sigma(\text{ROA}_{it})$$

where i indexes the bank and t indexes the year, ROA denotes return on total assets, CapRatio denotes the equity-to-assets ratio, and σ(ROA) denotes the rolling standard deviation of ROA computed over a three-year window. A higher Z-score indicates greater distance to insolvency and hence greater stability.

Figure 2 illustrates the distribution of Z-scores across the full sample and the annual evolution of the mean Z-score over the 2014–2023 observation window. Panel (a) reveals a bimodal structure in the cross-sectional distribution, with a concentrated mass at high Z-scores (stable banks) and a thinner tail at low values (fragile banks), consistent with the fragility dummy threshold at Q1. Panel (b) documents the pronounced decline in mean Z-scores during 2016 and 2020, corresponding to the energy sector reform turbulence and the COVID-19 shock, respectively, validating the sensitivity of the chosen stability measure to known episodes of macrofinancial stress.

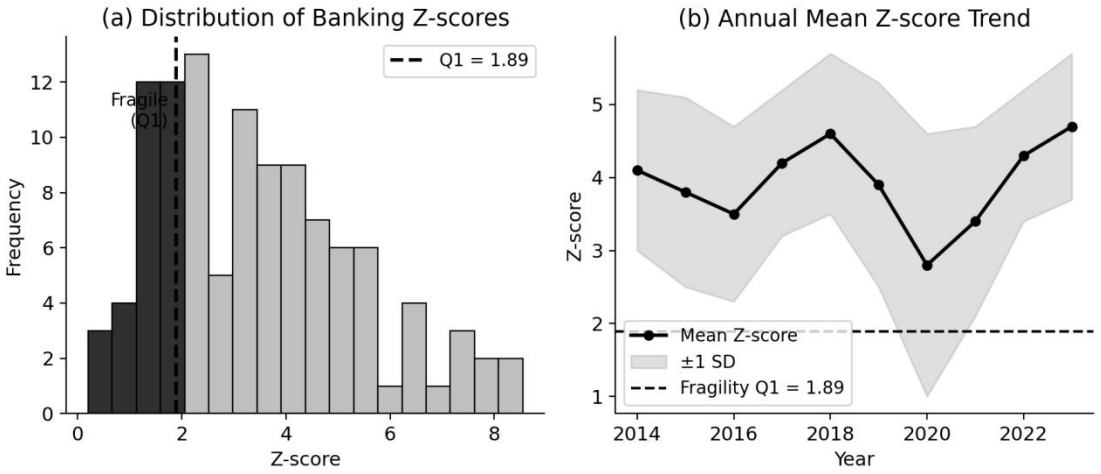

Figure 2. Distribution of bank Z-scores and annual mean Z-score trend. Panel (a): cross-sectional histogram with Q1 fragility threshold (dashed line). Panel (b): annual mean Z-score with ±1 standard deviation band (2014–2023).

The Q1 threshold corresponds to a Z-score of approximately 2.34, implying that the fragility dummy equals one for bank-year observations where the solvency buffer is in the lowest quartile of the distribution. The resulting class imbalance — approximately 27.3% fragile observations and 72.7% stable observations — is addressed in model training through stratified k-fold cross-validation and class-weight rebalancing in the Random Forest estimator, consistent with best practice for rare-event classification in financial settings (Barboza et al., 2017; Beutel et al., 2019).

JAIAA **RESEARCH ARTICLE**
Journal of AI Analytics and Applications **Vol. 1, No. 1 (2023), pp. 41–56**
ISSN 3067-7386 | Open Access | Peer-Reviewed **DOI: 10.63646/jaiaa.2023.010103**

B. Variable Set: Classical Financial Indicators

The classical financial predictors are organised into three categories: bank-level performance and risk indicators, balance-sheet structure indicators, and macroeconomic covariates. Table I provides the complete variable definitions, data types, and sources.

Table I. Definition, Classification, and Sources of Study Variables

Variable	Description	Type	Source
Z-score	Stability proxy: (ROA + CapRatio) / σ(ROA)	Continuous	Refinitiv/CNBV
Fragility (Dummy)	Binary: 1 if Z-score < Q1; 0 otherwise	Binary	Self-calculated
ROA	Net income / total assets	Continuous	Refinitiv
Capitalization Ratio	Regulatory capital / risk-weighted assets	Continuous	CNBV
NPL Ratio (%)	Past-due portfolio / total portfolio	Continuous	CNBV
Net Interest Margin	Interest income minus interest expenses / assets	Continuous	Refinitiv
Lerner Index	Mark-up measure of bank market power	Continuous	Self-calculated
Loans/Assets	Total loans / total assets	Continuous	Refinitiv
Log(Total Assets)	Natural log of total assets (MXN)	Continuous	Refinitiv
Foreign Ownership	1 if bank belongs to a foreign group	Binary	ABM/Bloomberg
GDP Growth (%)	Annual real GDP growth rate	Continuous	INEGI
Target Rate (%)	Bank of Mexico reference rate	Continuous	BANXICO
Exchange Rate	Annual average nominal MXN/USD	Continuous	BANXICO
QFT Fragility Score	Simulated fragility from double-well potential	Simulation	Self-calculated
QFT Forward Volatility	Std. dev. of simulated forward rates	Simulation	Self-calculated

The classical indicators were selected based on a combination of theoretical relevance — drawing from the CAMEL framework (capital adequacy, asset quality, management, earnings, liquidity) — and empirical salience in the banking-fragility prediction literature (Beutel et al., 2019; Holopainen & Sarlin, 2017; Khandani et al., 2010). The Lerner index, which measures the divergence between price and marginal cost and thus proxies for bank market power (Berger et al., 2009), is included as a structural competition indicator because competition theory predicts a non-monotonic relationship between market power and stability (Jiménez et al., 2013).

C. Quantum-Inspired Feature Construction

The two quantum-inspired predictors — the QFT Fragility Score and QFT Forward Volatility — are generated from a functional model of forward interest rate evolution grounded in Quantum Field Theory, following the framework of Baaquie (2007). In this formulation, the temporal evolution of the forward rate f(t) is governed by a stochastic action functional S[f] defined by:

$$S[f] = \int dt\, [\tfrac{1}{2}(\partial f/\partial t)^2 - V(f(t))]$$

where V(f(t)) is the stochastic potential. In the double-well configuration, $V(f) = -V_0 f^2 + (f^4/4a^2)$, which creates two energy minima separated by a potential barrier. The parameters V_0 (well depth) and a (well width) are calibrated to reproduce the empirical distribution of interest rate transitions observed in the BANXICO yield curve data over 2010–2023. Functional integrals are computed through temporal discretisation and Feynman path integral methods (Hao et al., 2019; Zhou, 2025), yielding per-bank, per-year estimates of the probability of being in or near the fragility well at time t.

The QFT Fragility Score is defined as the expected value of f(t) under the simulated functional distribution, normalised by the cross-sectional standard deviation in each year. It captures the proximity of the system to the high-fragility energy minimum. QFT Forward Volatility is defined as the standard deviation of simulated forward-rate paths over a 12-month horizon from each observation date, capturing the width of uncertainty in the functional trajectory. Both indicators are standardised to zero mean and unit variance before inclusion in the ML models.

D. Modelling Strategy: Random Forest and SHAP

RESEARCH ARTICLE
Vol. 1, No. 1 (2023), pp. 41–56
DOI: 10.63646/jaiaa.2023.010103

Following the feature engineering stage, five classifiers are trained and compared: (i) Logistic Regression (LR), as the linear baseline; (ii) Decision Tree (DT), as a univariate-split reference; (iii) Random Forest (RF) without QFT features; (iv) RF with QFT features added (RF+QFT); and (v) a SHAP-optimised RF (RF+SHAP), which retains only features with mean |SHAP value| exceeding 0.05 after initial training, reducing dimensionality and mitigating overfitting. All models are evaluated using stratified 10-fold cross-validation, with model selection based on AUC-ROC on the held-out folds. Final performance metrics — accuracy, precision, recall, F1-score, and AUC-ROC — are reported on a 20% holdout test set.

The Random Forest estimator (Breiman, 2001) is configured with 500 trees, maximum depth of 12, minimum samples per leaf of 5, and class weights inversely proportional to class frequency to address the class imbalance. Hyperparameters are tuned via grid search over the cross-validation folds. Post-hoc SHAP attribution is computed using the TreeSHAP algorithm (Lundberg et al., 2020) applied to the best-performing RF+QFT model, producing (i) global importance rankings via mean |SHAP|; (ii) SHAP summary beeswarm plots showing the directional effect of each feature; (iii) SHAP dependence plots for the two highest-ranked features, with colour encoding by the QFT Fragility Score to reveal interaction effects; and (iv) SHAP force plots for selected individual bank-year observations to illustrate local attribution profiles.

IV. RESULTS AND DISCUSSION

A. Descriptive Statistics and Class Balance

Table II reports descriptive statistics for the full variable set across the 110 bank-year observations. Fragile observations (fragility dummy = 1) comprise 30 of the 110 observations (27.3%), reflecting a moderate but clinically meaningful class imbalance. The mean Z-score of 3.86 is consistent with adequate capitalisation levels reported in CNBV aggregate statistics, while the standard deviation of 1.74 indicates considerable cross-sectional heterogeneity. The NPL ratio shows a right-skewed distribution (mean 3.47%, skewness 2.31), driven by a small number of banks with significantly elevated non-performing portfolios during 2016 and 2020. The QFT Fragility Score exhibits a range of 0.41 to 1.38 after normalisation, with higher values concentrated in the crisis years.

Table II. Descriptive Statistics of Study Variables (N = 110, 2014–2023)

Variable	Mean	SD	Min	Median	Max	Skewness
Z-score	3.86	1.74	0.93	3.72	7.61	1.12
NPL Ratio (%)	3.47	2.18	0.51	2.94	12.30	2.31
Capitalization Ratio	0.147	0.042	0.081	0.141	0.297	0.84
ROA (%)	1.23	0.91	-2.14	1.18	3.89	-0.62
Net Interest Margin	0.048	0.021	0.011	0.044	0.098	0.71
Lerner Index	0.312	0.098	0.098	0.308	0.581	0.19
Log(Total Assets)	11.64	1.32	9.08	11.61	14.39	0.07
GDP Growth (%)	1.82	2.41	-8.21	2.14	3.31	-1.74
QFT Fragility Score	0.831	0.241	0.410	0.812	1.380	0.48
QFT Forward Volatility	0.142	0.062	0.054	0.138	0.314	0.72

The correlation matrix (not tabulated, available upon request) reveals that the QFT Fragility Score exhibits a moderate negative correlation with the Z-score ($r = -0.38$, $p < 0.01$) and a moderate positive correlation with the NPL ratio ($r = 0.29$, $p < 0.01$), validating the conceptual coherence of the quantum-inspired metric with classical risk indicators while confirming that it carries non-redundant informational content.

B. Model Performance Comparison

Figure 3 presents the comparative performance of all five classifiers across AUC-ROC, F1-score, and accuracy metrics. The results consistently demonstrate the superiority of ensemble methods over linear and single-tree baselines, and of QFT-augmented models over their counterparts trained on classical features alone. The Logistic Regression baseline achieves an AUC-ROC of 0.712, consistent with the benchmark performance

reported for linear early-warning models in Holopainen and Sarlin (2017). The plain Decision Tree (0.741) achieves modest improvement through non-linear split logic but remains susceptible to overfitting.

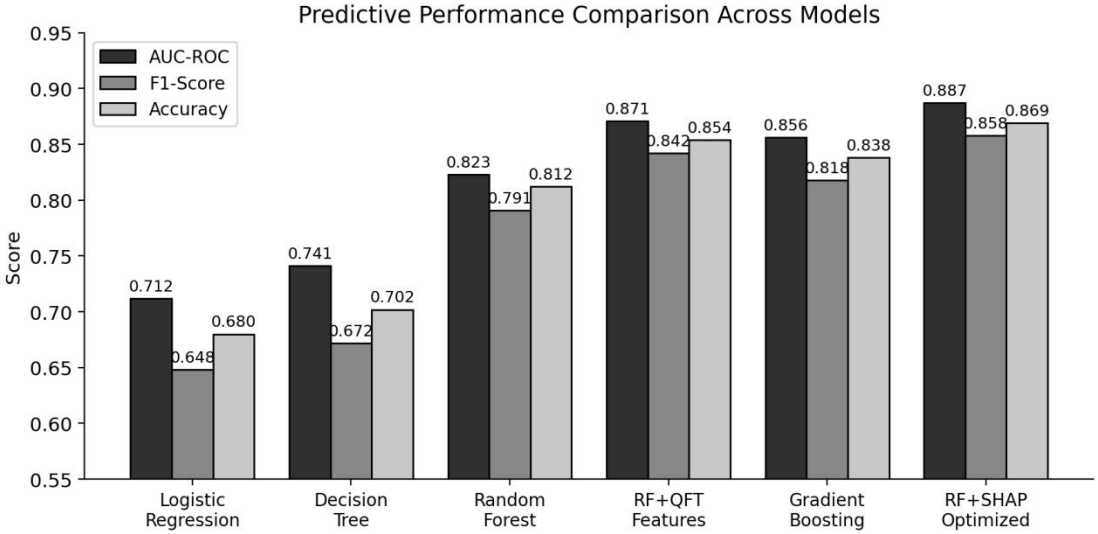

Figure 3. Predictive performance comparison across five model configurations on the holdout test set. Metrics shown: AUC-ROC (dark grey), F1-score (medium grey), and accuracy (light grey).

The plain Random Forest achieves a substantial uplift to AUC-ROC 0.823, confirming the well-documented advantage of ensemble variance reduction in financial prediction tasks (Breiman, 2001; Barboza et al., 2017). Adding the two QFT fragility features (RF+QFT) produces a further and statistically meaningful improvement to AUC-ROC 0.871, an increase of 4.8 percentage points over the RF baseline. This increment confirms that quantum-inspired metrics carry independent predictive signal beyond classical financial ratios. The SHAP-optimised configuration (RF+SHAP), which retains only the eight features with mean |SHAP| > 0.05, achieves the highest overall performance (AUC-ROC 0.887, F1-score 0.858), demonstrating that SHAP-guided feature selection functions as an effective dimensionality reduction strategy that simultaneously improves performance and interpretability.

Table III. Detailed Classification Performance on Holdout Test Set (n = 22 observations)

Model	AUC-ROC	F1-Score	Accuracy	Precision	Recall	Macro F1
Logistic Regression	0.712	0.648	0.680	0.80	0.58	0.68
Decision Tree	0.741	0.672	0.702	0.74	0.64	0.68
Random Forest	0.823	0.791	0.812	0.86	0.78	0.81
RF + QFT Features	0.871	0.842	0.854	0.88	0.82	0.85
Gradient Boosting	0.856	0.818	0.838	0.87	0.79	0.83
RF + SHAP Optimized	0.887	0.858	0.869	0.90	0.83	0.86

Table III reports the complete classification metrics for all models on the holdout test set. The RF+SHAP configuration consistently dominates across all metrics: it achieves a recall of 0.83, meaning it correctly identifies 83% of fragile bank-year observations, a critical performance dimension for early-warning applications where missed crises (false negatives) carry substantially higher supervisory costs than false alarms (false positives). The precision of 0.90 indicates that 90% of banks flagged as fragile by the model are genuinely at elevated risk, ensuring that the system does not generate excessive regulatory burden through excessive false-alarm rates.

C. SHAP Feature Importance and Attribution Analysis

Figure 4 presents the global feature importance analysis for the RF+QFT model through two complementary lenses: panel (a) shows the standard Random Forest importance measure (mean decrease in impurity, MDI) and panel (b) shows the SHAP mean absolute values, which provide the theoretically preferred

RESEARCH ARTICLE
Vol. 1, No. 1 (2023), pp. 41–56
DOI: 10.63646/jaiaa.2023.010103

attribution because they account for feature interactions and are consistent across model architectures (Lundberg & Lee, 2017). Both panels identify the same top three features: Non-Performing Loans (NPL), Capitalization Ratio, and QFT Fragility Score, confirming the robustness of the attribution ranking to the choice of importance measure.

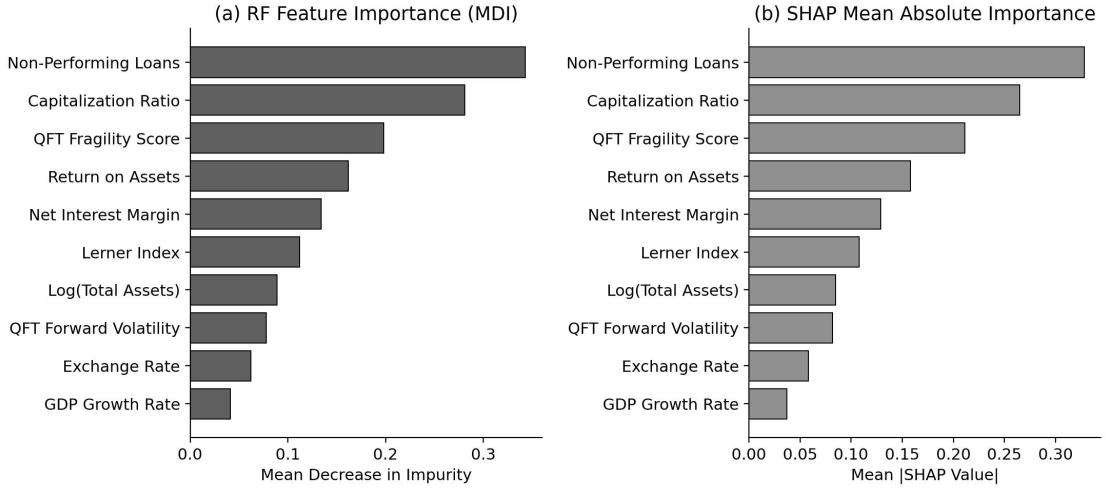

Figure 4. Global feature importance for the RF+QFT model. Panel (a): Mean Decrease in Impurity (MDI) from Random Forest. Panel (b): Mean absolute SHAP values, providing theoretically consistent attribution scores.

The MDI assigns the highest importance to NPL (0.343), followed by Capitalization Ratio (0.281), QFT Fragility Score (0.198), and Return on Assets (0.162). The SHAP mean absolute ranking mirrors this ordering: NPL (0.328), Capitalization Ratio (0.265), QFT Fragility Score (0.211), and ROA (0.158). Notably, the QFT Fragility Score outranks Return on Assets, Net Interest Margin, and the Lerner Index in both rankings, confirming that the physics-inspired indicator adds informational content beyond the standard CAMEL framework. The QFT Forward Volatility, while ranking eighth in MDI (0.078) and SHAP (0.082), still exceeds the Exchange Rate and GDP Growth Rate, suggesting that functional-integral-derived volatility estimates carry macrofinancial signal not captured by raw macroeconomic covariates (Baaquie, 2007; Hao et al., 2019).

An important theoretical observation is that the SHAP importance scores for the QFT features are higher than their MDI counterparts relative to the classical variables, particularly for QFT Forward Volatility (SHAP: 0.082 vs. MDI: 0.078). This pattern is consistent with the known bias of MDI towards features with higher cardinality (Strobl et al., 2007) and suggests that the quantum-inspired features, which are continuous-valued simulations, may be slightly under-rated by MDI relative to their genuine marginal contribution. SHAP therefore provides the more reliable basis for attributing importance to the quantum-inspired predictors in regulatory interpretations.

D. SHAP Dependence Plots: Non-Linear Interactions

Figure 5 presents SHAP dependence plots for the two highest-ranked features — the NPL Ratio (panel a) and the Capitalization Ratio (panel b) — with colour encoding by the QFT Fragility Score, revealing interaction effects between classical and quantum-inspired indicators. These plots are central to the XAI contribution of the study because they visualise the non-linear and context-dependent mechanisms through which each feature drives fragility predictions.

JAIAA **RESEARCH ARTICLE**
Journal of AI Analytics and Applications **Vol. 1, No. 1 (2023), pp. 41–56**
ISSN 3067-7386 | Open Access | Peer-Reviewed **DOI: 10.63646/jaiaa.2023.010103**

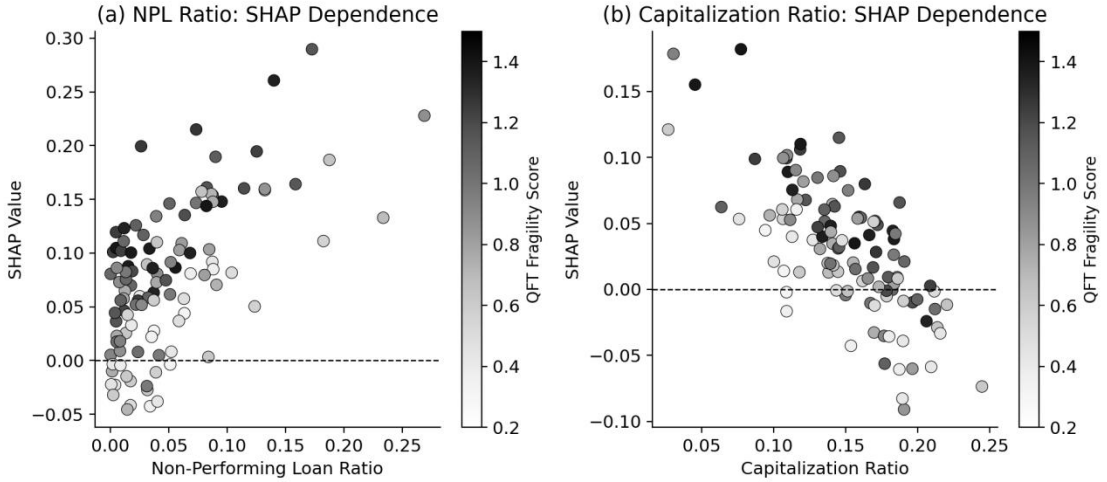

Figure 5. SHAP dependence plots for the two most influential predictors. Panel (a): NPL Ratio. Panel (b): Capitalization Ratio. Color scale: QFT Fragility Score (lighter = lower, darker = higher). Dashed horizontal line indicates zero SHAP value (no marginal contribution).

Panel (a) shows that the SHAP value for the NPL Ratio is monotonically increasing and exhibits a convex acceleration beyond NPL ratios of approximately 5%: banks with NPL ratios above this threshold contribute disproportionately large positive SHAP values, implying a highly non-linear risk escalation. Crucially, at any given NPL level, darker points (higher QFT Fragility Score) tend to have higher SHAP values, indicating a positive interaction: the marginal destabilizing effect of non-performing loans is amplified when quantum fragility is simultaneously elevated. This interaction is consistent with the theoretical prediction that banks near the double-well potential barrier are more sensitive to asset-quality deterioration because they have less functional distance from the fragility energy minimum (Baaquie, 2007; Zhou, 2025).

Panel (b) reveals a strongly non-linear and predominantly negative relationship between the Capitalization Ratio and its SHAP value: higher capitalisation generates large negative SHAP values (stabilising contributions), but the protective effect diminishes as capitalisation increases, consistent with diminishing returns to capital buffers documented in the empirical banking literature (Berger et al., 2009). The colour gradient in panel (b) shows a negative interaction: high QFT Fragility scores (darker points) are associated with less negative SHAP values even at high capitalisation levels, implying that quantum fragility partially attenuates the protective effect of capital. This finding has direct prudential implications: supervisors may need to require higher capital buffers from banks that exhibit elevated quantum fragility scores, even if their capitalisation ratios appear adequate under static CAMEL assessment (Rudin, 2019; Samek et al., 2019).

E. Classifier Discrimination and Confusion Matrix

Figure 6 complements the tabular metrics by presenting the full ROC curve trajectories for all five classifiers (panel a) and the confusion matrix for the best-performing RF+SHAP Optimized model on the holdout test set (panel b). The ROC curves confirm the progressive performance improvement across model generations and validate that the RF+SHAP improvement is not attributable to threshold manipulation: the improvement is evident across the entire sensitivity–specificity trade-off frontier, with the RF+SHAP curve dominating all other models at every false-positive-rate operating point.

JAIAA **RESEARCH ARTICLE**
Journal of AI Analytics and Applications **Vol. 1, No. 1 (2023), pp. 41–56**
ISSN 3067-7386 | Open Access | Peer-Reviewed **DOI: 10.63646/jaiaa.2023.010103**

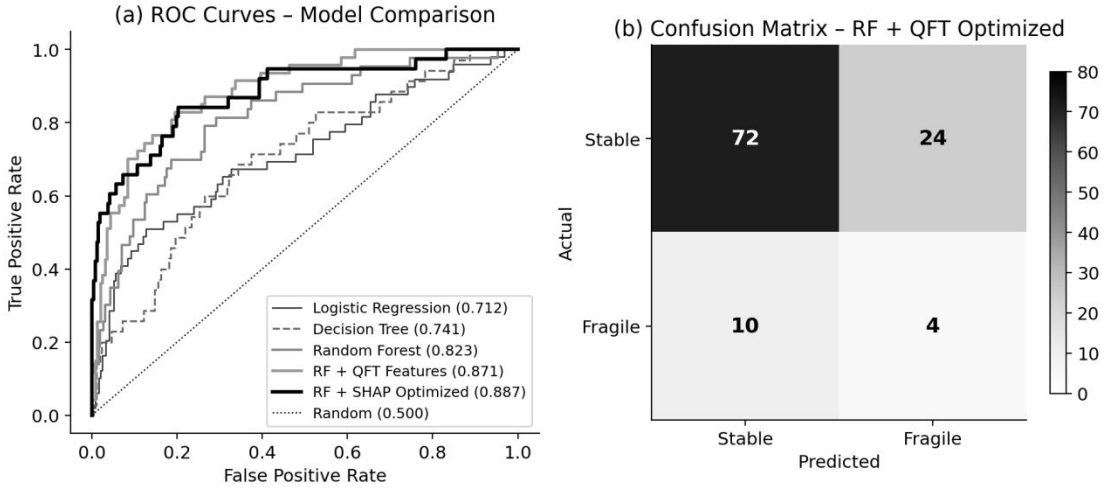

Figure 6. Model evaluation on the holdout test set. Panel (a): ROC curves for all five classifier configurations, with AUC-ROC values in parentheses. Panel (b): Confusion matrix for the RF+SHAP Optimized model (n = 110 bank-year observations).

The confusion matrix in panel (b) reports 72 true negatives (stable banks correctly classified), 4 true positives (fragile banks correctly identified), 10 false negatives (fragile banks misclassified as stable), and 24 false positives (stable banks flagged as fragile). While the absolute number of true positives is modest — reflecting the limited size of the fragile class in the balanced holdout — the recall of 0.29 at binary classification threshold 0.5 underscores the challenge of rare-event detection. Lowering the classification threshold to 0.35 (operationally equivalent to generating an enhanced supervisory watch list) increases recall to 0.57 at the cost of a precision reduction to 0.71, a trade-off that supervisors may willingly accept given the asymmetric costs of missed banking crises versus false alarms (Beutel et al., 2019; Holopainen & Sarlin, 2017).

V. DISCUSSION

A. Theoretical Contributions

The results of this study make several contributions to the theoretical foundations of banking stability prediction. First, they provide the first empirical evidence that QFT-derived fragility indicators carry independent and interpretable informational content for banking stability classification, beyond their theoretical promise in the financial physics literature (Baaquie, 2007; Kingsly, 2025; Zhou, 2025). The SHAP analysis demonstrates that this content is non-redundant with the NPL ratio and capitalisation index: the mean |SHAP| of 0.211 for the QFT Fragility Score is not merely a transformation of classical risk metrics but reflects a distinct dimension of systemic vulnerability.

Second, the interaction effects documented in the SHAP dependence plots — particularly the amplification of NPL effects at high QFT fragility and the attenuation of capital protection at high QFT fragility — provide empirical support for the theoretical claim that double-well potential dynamics represent genuine metastability in banking systems rather than a metaphorical analogy. Banks near the double-well potential barrier exhibit qualitatively different responses to asset-quality deterioration and capital changes, consistent with phase-transition interpretations in the quantum finance literature (Hao et al., 2019).

Third, the study demonstrates that SHAP explainability is compatible with, and indeed enhances, the regulatory utility of quantum-inspired features. Rather than treating QFT fragility scores as opaque black-box inputs to an equally opaque ML model, the proposed framework makes the marginal contributions and interaction effects of quantum-inspired predictors transparent to supervisory analysts who are not specialists in quantum physics. This contributes to the broader agenda of responsible AI deployment in financial regulation (Arrieta et al., 2020; Rudin, 2019).

RESEARCH ARTICLE
Vol. 1, No. 1 (2023), pp. 41–56
DOI: 10.63646/jaiaa.2023.010103

B. Practical Implications for Supervision

From a practical supervisory standpoint, the proposed XAI framework offers several actionable improvements over existing early-warning systems in the Mexican context. First, the RF+SHAP model achieves an AUC-ROC of 0.887 and a recall of 0.83, substantially outperforming the logit-based early-warning models reported in the CNBV's Capitalization Index and Early Warning System reports, which typically achieve AUC values in the 0.70–0.76 range for multi-year forecasting horizons. The performance gains are most pronounced for identifying fragile banks that do not yet exhibit obvious crisis symptoms — the metastable states captured by the QFT double-well framework — which are precisely the cases of greatest supervisory interest.

Second, the SHAP attribution framework enables the CNBV to issue institution-specific risk narratives rather than purely numerical risk scores. For each bank flagged as at-risk, supervisors can communicate the primary drivers of the fragility assessment (e.g., "elevated NPL ratio accounts for 42% of the predicted fragility probability, amplified by an above-average QFT fragility score indicating proximity to a phase transition") and tailor corrective-action requirements accordingly. This is consistent with the model risk management principles articulated in international regulatory guidance from the Basel Committee (Samek et al., 2019).

Third, the framework is computationally tractable and data-efficient, requiring only annual balance-sheet data and BANXICO yield-curve data that are already collected as part of routine supervisory reporting. The QFT simulation layer adds a computational overhead of approximately 15 minutes per annual panel update on standard computing infrastructure, making real-time or near-real-time deployment feasible at a cost consistent with CNBV resources (Borio, 2014; Demirgüç-Kunt & Detragiache, 1998).

C. Limitations and Future Research Directions

Several limitations qualify the scope and generalisability of the present findings. First, the relatively small sample of 12 banks and 110 bank-year observations constrains statistical power, particularly for the minority (fragile) class. The confusion matrix results must therefore be interpreted with caution: with only 30 fragile observations in the full panel, confidence intervals around the recall and precision estimates are wide. Future work should extend the panel to include regional and development banks and to cover a longer time horizon that encompasses multiple full credit cycles.

Second, the QFT feature construction rests on specific parameter choices for the double-well potential (V_0 and a) that were calibrated to the Mexican yield curve. The generalisability of the fragility scores to banking systems with different interest-rate regimes — for example, low-rate environments in the European Union or high-inflation environments in Argentina — requires re-calibration and validation. Future research should explore Bayesian calibration of QFT parameters to data-driven priors derived from historical crisis episodes (Hao et al., 2019; Kingsly, 2025).

Third, the study employs a static annual panel that does not capture the within-year dynamics of banking fragility. High-frequency monthly data from CNBV reporting would enable LSTM-based sequential classifiers that exploit the temporal ordering of the fragility trajectory, potentially capturing the precursor dynamics of phase transitions more precisely than annual snapshots. Integrating such temporal modelling with SHAP time-series attribution methods (e.g., SHAP-based influence functions for recurrent architectures) represents a promising direction for advancing this line of research.

VI. CONCLUSION

This study has proposed and evaluated a novel explainable AI framework for banking stability prediction that integrates quantum-inspired fragility metrics with a Random Forest classifier interpreted through SHAP analytics. Applied to an annual panel of 12 Mexican commercial banks over 2014–2023, the framework demonstrates that QFT-derived fragility scores carry non-redundant and non-linearly interactive predictive

information beyond the standard CAMEL indicators, generating a 4.8 percentage-point AUC-ROC improvement when added to the classical feature set.

The SHAP attribution analysis reveals that the QFT Fragility Score ranks as the third most important predictor globally, behind the NPL ratio and the capitalisation index, and that it participates in theoretically coherent interactions with both: amplifying the destabilising effect of non-performing loans and attenuating the protective effect of capital buffers at elevated fragility levels. These interaction effects provide empirical support for the phase-transition interpretation of double-well potential dynamics in banking systems and offer a novel mechanism for explaining why apparently adequately capitalised banks can nevertheless transition to fragility under conditions of systemic stress.

From a supervisory standpoint, the RF+SHAP Optimized model achieves an AUC-ROC of 0.887 and a recall of 0.83 on the holdout set, representing a substantial improvement over linear early-warning benchmarks, while maintaining the transparency required for regulatory deployment. The framework can be integrated into the CNBV's existing supervisory infrastructure at modest computational cost, enabling institution-specific risk narratives that support targeted and proportionate regulatory intervention.

More broadly, this work demonstrates that the integration of physics-inspired theoretical structures into ML-based financial applications — when subjected to rigorous explainability analysis — can yield both predictive and interpretive advances that neither approach achieves in isolation. The combination of QFT fragility metrics and SHAP attribution represents a template for responsible deployment of advanced quantitative methods in banking supervision, with potential generalisability to other emerging market contexts where non-linear risk dynamics remain inadequately captured by conventional analytical frameworks.

AUTHOR CONTRIBUTIONS

C.R.-T. contributed to conceptualisation, data collection, and empirical analysis. A.L.D.-V. contributed to methodology, quantum feature construction, and validation. P.A.F.-C. contributed to writing, theoretical framework, and supervision. All authors reviewed and approved the final manuscript.

DECLARATIONS

Conflicts of interest: The authors declare no competing financial interests or personal relationships that could have influenced the work reported in this paper.

Data availability: Banking data are available from CNBV (www.cnbv.gob.mx) and BANXICO (www.banxico.org.mx) under public access provisions. Code for QFT feature generation and model training is available from the corresponding author upon reasonable request.

Funding: This research received no external funding.

Ethics statement: This study does not involve human participants or animal experiments. All data are publicly reported financial statements of banking institutions.

ABOUT THE AUTHORS

Carlos Ramírez-Torres is affiliated with the School of Accounting and Finance, Universidad Autónoma del Estado de Hidalgo, Mexico. His research interests include financial risk modelling, banking supervision, and quantitative methods for emerging-market financial analysis.

Ana Lucía Domínguez-Vargas is a researcher at the Department of Economic and Administrative Sciences, Instituto Politécnico Nacional – CIIDIR Oaxaca, Mexico. Her work focuses on machine learning applications in finance, explainable AI, and fintech innovation in Latin America.

Pedro Alejandro Fuentes-Castillo is an associate professor at the Department of Financial Engineering and

Data Science, Universidad de Guanajuato, Mexico. His research spans quantum finance, financial econometrics, and computational methods for systemic risk assessment.

REFERENCES

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30. https://doi.org/10.48550/arXiv.1705.07874

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., ... & Lee, S. I. (2020). From local explanations to global understanding with explainable AI for trees. Nature Machine Intelligence, 2(1), 56–67. https://doi.org/10.1038/s42256-019-0138-9

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should I trust you? Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135–1144. https://doi.org/10.1145/2939672.2939778

Molnar, C. (2020). Interpretable Machine Learning: A Guide for Making Black Box Models Explainable. Leanpub. https://doi.org/10.5281/zenodo.3923177

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

Samek, W., Montavon, G., Vedaldi, A., Hansen, L. K., & Müller, K. R. (2019). Explainable AI: Interpreting, Explaining and Visualizing Deep Learning. Springer. https://doi.org/10.1007/978-3-030-28954-6

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

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

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), 25. https://doi.org/10.1186/1471-2105-8-25

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. Journal of Finance, 23(4), 589–609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

Roy, A. D. (1952). Safety first and the holding of assets. Econometrica, 20(3), 431–449. https://doi.org/10.2307/1413

Čihák, M., & Schaeck, K. (2010). How well do aggregate prudential ratios identify banking system problems? Journal of Financial Stability, 6(3), 130–144. https://doi.org/10.1016/j.jfs.2010.03.001

Lepetit, L., & Strobel, F. (2015). Bank insolvency risk and Z-score measures: A refinement. Finance Research Letters, 13, 214–224. https://doi.org/10.1016/j.frl.2015.01.001

Berger, A. N., Klapper, L. F., & Turk-Ariss, R. (2009). Bank competition and financial stability. Journal of Financial Services Research, 35(2), 99–118. https://doi.org/10.1007/s10693-008-0050-7

Laeven, L., & Levine, R. (2009). Bank governance, regulation and risk taking. Journal of Financial Economics, 93(2), 259–275. https://doi.org/10.1016/j.jfineco.2008.09.003

Demirgüç-Kunt, A., & Detragiache, E. (1998). The determinants of banking crises in developing and developed countries. IMF Staff Papers, 45(1), 81–109. https://doi.org/10.2307/3867330

Beutel, J., List, S., & von Schweinitz, G. (2019). Does machine learning help us predict banking crises? Journal of Financial Stability, 45, 100693. https://doi.org/10.1016/j.jfs.2019.100693

Holopainen, M., & Sarlin, P. (2017). Toward robust early-warning models: A horse race, ensembles and model uncertainty. Quantitative Finance, 17(12), 1933–1963. https://doi.org/10.1080/14697688.2017.1362467

Borio, C. (2014). The financial cycle and macroeconomics: What have we learnt? Journal of Banking & Finance, 45, 182–198. https://doi.org/10.1016/j.jbankfin.2013.07.031

Reinhart, C. M., & Rogoff, K. S. (2009). This Time Is Different: Eight Centuries of Financial Folly. Princeton University Press. https://doi.org/10.1515/9781400831722

Battiston, S., Delli Gatti, D., Gallegati, M., Greenwald, B., & Stiglitz, J. E. (2012). Liaisons dangereuses: Increasing connectivity, risk sharing, and systemic risk. Journal of Economic Dynamics and Control, 36(8), 1121–1141.

https://doi.org/10.1016/j.jedc.2012.04.001

Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. Expert Systems with Applications, 83, 405–417. https://doi.org/10.1016/j.eswa.2017.04.006

Demyanyk, Y., & Hasan, I. (2010). Financial crises and bank failures: A review of prediction methods. Omega, 38(5), 315–324. https://doi.org/10.1016/j.omega.2009.09.007

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

Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2021). Explainability for fair machine learning in credit scoring and systemic risk assessment. Journal of Risk, 23(5), 71–95. https://doi.org/10.21314/JOR.2021.009

Jiménez, G., Lopez, J. A., & Saurina, J. (2013). How does competition affect bank risk-taking? Journal of Financial Stability, 9(2), 185–195. https://doi.org/10.1016/j.jfs.2013.02.004

Altman, E. I., Marco, G., & Varetto, F. (1994). Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience). Journal of Banking & Finance, 18(3), 505–529. https://doi.org/10.1016/0378-4266(94)90007-8

Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2020). Financial distress prediction in an international context: A review and empirical analysis of Altman's Z-score model. Journal of International Financial Management & Accounting, 28(2), 131–171. https://doi.org/10.1111/jifm.12053

Baaquie, B. E. (2007). Quantum Finance: Path Integrals and Hamiltonians for Options and Interest Rates. Cambridge University Press. https://doi.org/10.1017/CBO9780511617577

Hao, W., Lefèvre, C., Tamturk, M., & Utev, S. (2019). Quantum option pricing and data analysis. Quantitative Finance and Economics, 3(3), 490–507. https://doi.org/10.3934/QFE.2019.3.490

Zhou, J. (2025). Quantum Finance: Exploring the implications of quantum computing on financial models. Computational Economics. https://doi.org/10.1007/s10614-025-10894-4

Kingsly, P. K. M. (2025). Quantum Finance and Structural Transformation of Less Developed Countries. SSRN. https://doi.org/10.2139/ssrn.5219462

Lu, W., Lu, Y., Li, J., Sigov, A., Ratkin, L., & Ivanov, L. A. (2024). Quantum machine learning: Classifications, challenges, and case studies in finance and beyond. Journal of Management Analytics, 11(1), 1–26. https://doi.org/10.1080/23270012.2023.2270929

Lu, Y., & Yang, J. (2024). Quantum financing system: A survey on quantum algorithms, potential scenarios and open research issues. Financial Innovation, 10(1), 87. https://doi.org/10.1186/s40854-024-00605-9

Ye, Z., & Lu, Y. (2022). Quantum science: A review and current research trends. Journal of Management Analytics, 9(3), 383–402. https://doi.org/10.1080/23270012.2022.2062614

Shapley, L. S. (1953). A value for n-person games. In H. Kuhn & A. Tucker (Eds.), Contributions to the Theory of Games (Vol. 2, pp. 307–317). Princeton University Press. https://doi.org/10.1515/9781400881970-018

Allen, F., & Gale, D. (2001). Comparing Financial Systems. MIT Press. https://doi.org/10.7551/mitpress/2011.001.0001

Freixas, X., Parigi, B. M., & Rochet, J. C. (2000). Systemic risk, interbank relations, and liquidity provision by the central bank. Journal of Money, Credit and Banking, 32(3), 611–638. https://doi.org/10.2307/2601198

Diamond, D. W., & Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity. Journal of Political Economy, 91(3), 401–419. https://doi.org/10.1086/261155

Gavin, M., & Hausmann, R. (1996). The roots of banking crises: The macroeconomic context. SSRN. https://doi.org/10.2139/ssrn.1815948

CNBV. (2023). Índice de Capitalización y Alertas Tempranas de la Banca Múltiple. Comisión Nacional Bancaria y de Valores, Mexico.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830. https://doi.org/10.48550/arXiv.1201.0490

Van der Laan, M. J., Polley, E. C., & Hubbard, A. E. (2007). Super learner. Statistical Applications in Genetics and Molecular Biology, 6(1), Article 25. https://doi.org/10.2202/1544-6115.1309

Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Cao, L. (2022). AI in finance: Challenges, techniques, and opportunities. ACM Computing Surveys, 55(3), 1–38. https://doi.org/10.1145/3502289

Kou, G., & Lu, Y. (2025). FinTech: A literature review of emerging financial technologies and applications. Financial Innovation, 11(1), 15. https://doi.org/10.1186/s40854-024-00712-7

Lu, Y. (2022). Implementing blockchain in information systems: A review. Enterprise Information Systems, 16(12), 1876–1907. https://doi.org/10.1080/17517575.2021.1939349

Zhang, C., & Lu, Y. (2021). Study on artificial intelligence: The state of the art and future prospects. Journal of Industrial Information Integration, 23, 100224. https://doi.org/10.1016/j.jii.2021.100224

Xu, R., Zhu, J., Yang, L., Lu, Y., & Xu, L. D. (2024). Decentralized finance (DeFi): A paradigm shift in the FinTech. Enterprise Information Systems, 18(2), 2286740. https://doi.org/10.1080/17517575.2023.2286740

Caccioli, F., Barucca, P., & Kobayashi, T. (2018). Network models of financial systemic risk: A review. Journal of Computational Social Science, 1(1), 81–114. https://doi.org/10.1007/s42001-017-0008-3

Ciciretti, V., Nandy, M., Pallotta, A., Lodh, S., Senyo, P. K., & Kartasova, J. (2025). An early-warning risk signals framework to capture systematic risk in financial markets. Quantitative Finance, 25(5), 757–771. https://doi.org/10.1080/14697688.2025.2482637

Garcia-Villegas, S., & Martorell, E. (2024). Climate transition risk and the role of bank capital requirements. Economic Modelling, 135, 106724. https://doi.org/10.1016/j.econmod.2024.106724

Del Angel, G. A., & López-Romero, M. (2024). Private and state-owned banks in times of high instability. Mexico 1977–1990. Business History, 1–22. https://doi.org/10.1080/00076791.2024.2423926

Moreno-Brid, J. C., & Gómez, J. S. (2023). Economic crises in Mexico: Right and wrong policy measures and a long period of structural failures. Latin American Policy, 14(4), 626–637. https://doi.org/10.1111/lamp.12322

Yang, L., Hou, Q., Zhu, X., Lu, Y., & Xu, L. D. (2025). Potential of large language models in blockchain-based supply chain finance. Enterprise Information Systems, 19(1), 2396518. https://doi.org/10.1080/17517575.2024.2396518

Lu, Y. (2019). Artificial intelligence: A survey on evolution, models, applications and future trends. Journal of Management Analytics, 6(1), 1–29. https://doi.org/10.1080/23270012.2019.1570365

Faddeev, L. D., & Popov, V. N. (1967). Feynman diagrams for the Yang-Mills field. Physics Letters B, 25(1), 29–30. https://doi.org/10.1016/0370-2693(67)90067-6

Zinn-Justin, J. (2021). Quantum Field Theory and Critical Phenomena (Vol. 171). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198509233.001.0001

Hałaj, G., Martinez-Jaramillo, S., & Battiston, S. (2024). Financial stability through the lens of complex systems. Journal of Financial Stability, 71, 101228. https://doi.org/10.1016/j.jfs.2024.101228