International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

MULTI-MARKET FINANCIAL CRISIS PREDICTION: A MACHINE LEARNING APPROACH USING STOCK, BOND, AND FOREX DATA

Rejon Kumar Ray[1], Proshanta Kumar Bhowmik[2], Badruddowza[3], Mahamuda Khanom[4], Redwan Ahmed Ratul[5], Md Toushif Pramanik[6], Sonia Afroze[7], Md Ekramul Hoque[8], Riad Hossain[9] and Kazi Md Shahadat Hossain[10]

[1]MBA in Business Analytics, Gannon University

[2]Department of Business Analytics, Trine University, Angola, IN, USA

[3]Master of Science in Computer and Information Systems, Gannon University, Erie, Pa

[4]Master of Science in Information Technology, Washington University of Science and Technology

[5]Applied Computing: Artificial Intelligence, University of Arizona

[6]Department of Electrical and Cyber Engineering, Gannon University, Erie, PA, USA

[7]MSc in Botany, Jahangirnagar University

[8]Master of Science in Business Analytics, Trine University, Angola, Indiana

[9]Master of Science in Computer Science, Central Michigan University, Michigan, USA

[10]Master of Business Administration in Logistics Management, Central Michigan University, Michigan, USA

Corresponding Author: Rejon Kumar Ray, **Email:** ray.rejon@gmail.com

Abstract

Early detection of financial crises remains a persistent challenge because traditional single-market indicators, focused separately on equities, bonds, or foreign exchange, cannot fully capture cross-market contagion and volatility spillovers that drive systemic breakdowns. This study develops a unified machine learning framework for multi-market crisis prediction in the United States by integrating data from the S&P 500 index, U.S. Treasury bond yields and prices, and EUR/USD exchange rates spanning 2010–2025. The pipeline constructs engineered features that encompass daily and rolling log-returns, multi-horizon volatilities, bond yield-curve slopes, contagion fractions, and pairwise rolling correlations, alongside principal component decompositions to capture latent systemic factors. Using rolling time-series cross-validation, we benchmark Logistic Regression and XGBoost models under realistic temporal splits and class imbalance adjustments. Results show that multi-market fusion markedly improves predictive power over single-market baselines, with XGBoost achieving superior ROC-AUC and precision–recall scores while maintaining calibration stability. SHAP-based explainability identifies volatility clustering, yield-slope inversions, and stock-bond correlation

Received: August 03,2025

spikes as dominant early-warning features. Economic backtests further demonstrate that a crisis-aware portfolio allocation strategy informed by model probabilities reduces maximum drawdown and enhances risk-adjusted returns relative to buy-and-hold benchmarks. Collectively, these findings establish that machine-learning-driven, cross-market integration provides a transparent and operationally feasible foundation for systemic-risk monitoring and financial-stability forecasting within U.S. markets.

Keywords: Systemic Risk, Financial Crisis Prediction, XGBoost, Logistic Regression, SHAP, Multi-Market Modeling, U.S. Treasury Bonds, S&P 500, EUR/USD, Time-Series Machine Learning

1. Introduction

1.1 Background and Motivation

Financial crises emerge not as isolated market failures but as cascading phenomena propagating through deeply interlinked asset classes, equities, fixed income, and foreign exchange, each transmitting volatility and liquidity shocks to the others through multiple contagion channels. Traditional early-warning systems have long relied on single-domain indicators such as credit spreads, stock market volatility, or yield curve inversions. However, these approaches systematically underrepresent the multidimensional nature of systemic instability. When markets become tightly coupled through risk-parity trades, global funding dependencies, and cross-asset hedging, local disturbances amplify into global dislocations. Gieck (2012) demonstrated that interdependence among asset markets increases significantly during stress episodes, implying that shocks transmitted via global portfolios and currency linkages can escalate into full-scale crises [13]. This insight underscores the insufficiency of monitoring isolated markets when systemic linkages are the very mechanism of contagion.

The limitations of single-market crisis predictors became particularly evident in the United States during recent episodes such as the 2020 COVID-19 pandemic crash and the 2022 interest rate shock. During these periods, volatility spikes in the S&P 500 coincided with abrupt yield-curve inversions and sharp U.S. dollar appreciation against the euro, collectively reflecting synchronized stress across equity, bond, and forex markets. Zhang et al. (2023) provided empirical confirmation that stock, bond, and foreign exchange volatilities exhibited heightened connectedness during COVID-19, amplifying systemic risk across asset classes [26]. Similarly, Bluwstein et al. (2020) found that combining credit growth and yield curve information improves predictive accuracy for financial crises relative to traditional univariate indicators, emphasizing the importance of integrating cross-market signals [3].

Machine learning methods have recently provided new tools to quantify and forecast such interconnected crises. Fouliard et al. (2021) argued that financial crises can be effectively anticipated by leveraging nonlinear learning methods that detect subtle early-warning features in macro-financial data, surpassing traditional econometric models in recall and lead time [11]. Chen et al. (2025) extended this insight by constructing a dynamic contagion network across global asset classes, showing that shifts in network centrality and clustering precede crisis phases [7]. Kohan Sefidi (2025) reached similar conclusions using multilayer network analysis

Received: August 03,2025

and LSTM-based spillover forecasting, demonstrating that learning architectures capturing intertemporal dependencies can identify crisis onset more reliably than static models [17]. Together, these findings reinforce the argument that market crises are inherently systemic, with nonlinear propagation patterns requiring models that jointly process multiple sources of financial data.

From a U.S. perspective, the increasing integration of equity, bond, and forex markets, amplified by algorithmic trading, macro-hedging, and global capital flows, necessitates a holistic modeling approach. Tölö (2020) emphasized that recurrent neural networks and other temporal learning architectures provide predictive advantage by learning from sequential inter-market dependencies rather than snapshot metrics [24]. The convergence of these empirical and methodological developments suggests that the next generation of crisis forecasting systems must explicitly model inter-market linkages. The inability to account for volatility spillovers, yield-curve distortions, and cross-currency capital shifts has repeatedly undermined early-warning frameworks, as seen in 2008, 2020, and 2022. Hence, a new paradigm is required, one that fuses heterogeneous data sources and interpretable learning architectures to uncover latent systemic stress patterns before they culminate in crisis.

1.2 Importance of This Research

The stability of U.S. financial markets is not merely a matter of sectoral health but of systemic coordination across asset classes that collectively anchor global capital flows. In this environment, the importance of a unified modeling framework integrating market signals from equities, bonds, and currencies cannot be overstated. Policymakers, institutional investors, and central banks increasingly depend on real-time systemic-risk diagnostics to anticipate contagion effects and deploy preemptive measures. Yet current models either lack granularity, focusing on a single market, or opacity, embedding predictive signals in black-box neural architectures devoid of interpretability. A transparent, data-driven early-warning system capable of explaining cross-market dynamics is therefore essential for informed policy and investment decisions. The integration of the S&P 500, U.S. Treasury yields, and the EUR/USD exchange rate provides a comprehensive lens for capturing three critical dimensions of financial stress. First, equity market volatility reflects investor sentiment and risk appetite; second, bond yield-curve dynamics encode expectations about monetary policy, inflation, and credit conditions; and third, foreign exchange movements signal international capital flows and external balance pressures. When modeled jointly, these markets reveal interaction patterns that single-domain indicators obscure. For instance, during tightening cycles, yield-curve inversions often precede equity selloffs as higher discount rates compress valuations, while simultaneous dollar appreciation accelerates capital outflows from risk assets, an alignment historically associated with recessionary onset.

Chen et al. (2025) provided empirical validation that these inter-market dependencies form a contagion network where changes in one domain rapidly alter the topology of others, influencing liquidity, volatility, and systemic centrality [6]. Similarly, Bluwstein et al. (2020) emphasized that yield-curve slopes serve as predictive markers of financial crises when combined with credit expansion metrics, reinforcing the need for cross-domain integration [31].

Received: August 03,2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Kohan Sefidi (2025) further demonstrated that machine learning models leveraging inter-market data can detect spillover effects between asset classes with superior timeliness and accuracy relative to traditional approaches [17]. These insights collectively point toward the necessity of unified modeling frameworks where systemic stress is represented not as a sequence of independent events but as a synchronized process of inter-market transmission. From a macroprudential perspective, such an approach holds clear benefits for regulatory institutions such as the Federal Reserve, the Office of Financial Research, and the Financial Stability Oversight Council, which must evaluate systemic vulnerabilities in near real time. The ability to decompose crisis risk into interpretable, data-driven signals enables not only more responsive policy interventions but also improved communication with market participants. For institutional investors, including hedge funds, pension funds, and sovereign wealth managers, these models support portfolio reallocation decisions under emerging stress conditions, improving resilience to tail events. As Gieck (2012) observed, understanding contagion dynamics allows policymakers to design countercyclical measures that dampen cross-market amplification mechanisms [13]. By combining insights from market microstructure, macroeconomics, and machine learning, a unified crisis-prediction system contributes to a more transparent and stable financial ecosystem.

1.3 Objectives an

d Contributions

The primary objective of this research is to develop a machine learning–based, multi-market framework for early crisis prediction in the United States that integrates signals from equity, bond, and forex markets into a single interpretable decision-support system. The study seeks to construct a feature space encompassing volatility measures, yield-curve slopes, and rolling cross-market correlations to quantify systemic interactions and detect contagion buildup. In doing so, it establishes the first comprehensive dataset combining U.S. stock, bond, and currency market data spanning 2010–2025, enabling robust temporal learning and historical backtesting across diverse macroeconomic regimes. A secondary objective is to evaluate the comparative performance of interpretable logistic regression and nonlinear XGBoost models within a time-aware rolling cross-validation framework, mitigating temporal leakage and ensuring realistic out-of-sample assessment. Model interpretability is further enhanced through SHAP analysis, allowing decomposition of crisis probabilities into underlying feature contributions and revealing how volatility clustering, yield-curve flattening, and equity-bond correlation surges drive predicted risk levels. The study also incorporates an economic backtesting component linking predicted crisis probabilities to simulated trading outcomes, demonstrating how early warnings translate into actionable investment strategies. The key contribution lies in bridging predictive performance with interpretability: combining the precision of ensemble learning with the transparency of regression-based signal attribution. By embedding explainable machine learning into systemic-risk forecasting, the proposed framework aligns financial stability monitoring with practical decision support for policymakers and investors alike.

2. Literature Review

Received: August 03,2025

International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

2.1 Traditional Financial Crisis Prediction

Traditional financial crisis prediction has long relied on macroeconomic and financial indicators, such as credit spreads, yield curve slopes, and volatility indices like the VIX, to signal market stress. These variables, when interpreted through classical econometric frameworks such as probit and logit models, have been used to estimate the likelihood of crisis events or recessions. Benzoni and Chyruk (2018) observed that the yield-curve slope, particularly the spread between long- and short-term U.S. Treasury rates, has historically served as a leading indicator of recessions because it encapsulates investor expectations about future economic conditions and monetary policy tightening [2]. However, the predictive power of such univariate indicators has diminished in the presence of complex, globalized markets where cross-asset interdependencies amplify systemic vulnerabilities. Econometric approaches such as vector autoregressions (VARs) and dynamic factor models have been widely used to capture intertemporal dependencies across macroeconomic variables, yet they struggle to generalize beyond stable historical relationships. Wang et al. (2023) highlighted that systemic risk predictability often deteriorates when macroeconomic regimes shift, as the underlying parameters of linear models are unable to adapt to nonlinear co-movements between markets during crisis periods [25]. These models tend to lag real-time developments, offering little foresight during periods of rapid contagion. Moreover, traditional indicators such as the VIX or credit spreads reflect market reaction to emerging shocks rather than their genesis, which limits their utility as proactive warning signals.

As financial systems became more globally integrated, early-warning frameworks rooted in econometric methods proved insufficient to capture cross-border spillovers and liquidity feedback loops. Casabianca et al. (2022) demonstrated that while macro-level stress indices and credit aggregates can rank crisis determinants effectively across countries, they exhibit weak temporal generalization due to evolving market structures and policy regimes [5]. The increasing speed of information propagation and algorithmic trading has further eroded the relevance of static linear models, whose lag structures fail to represent instantaneous feedback effects among equities, bonds, and currencies. In essence, traditional financial crisis prediction has been constrained by assumptions of linearity, stationarity, and independence between asset classes. The growing evidence of nonlinear contagion and dynamic feedback loops, particularly visible during the 2008 global financial crisis and the 2020 pandemic shock, demands more adaptive, data-driven methodologies capable of capturing hidden inter-market dependencies and real-time shifts in systemic risk. The literature increasingly recognizes that effective crisis forecasting must move beyond single-variable or static macro models toward integrative frameworks that leverage the full cross-section of market information.

2.2 Machine Learning in Financial Stability Forecasting

Machine learning has emerged as a transformative approach in financial stability forecasting, offering powerful tools to detect nonlinear dependencies and structural shifts that elude traditional econometric models. Casabianca et al. (2022) applied tree-based algorithms such as Random Forests and Gradient Boosting to identify key determinants of banking crises across countries, showing that nonlinear feature interactions substantially improve crisis prediction

Received: August 03,2025

International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

accuracy over conventional models [5]. Similarly, Chen et al. (2024) demonstrated that machine learning models, particularly XGBoost and Support Vector Machines, outperform traditional econometric benchmarks in forecasting stock market crash events by exploiting higher-order relationships among volatility, trading volume, and macro variables [7]. These advancements signal a paradigm shift toward data-driven systemic-risk monitoring. However, most studies in this domain focus on single-asset or single-sector applications, such as bank distress, corporate default, or stock market volatility. Tölö (2020) showed that recurrent neural networks (RNNs) can effectively predict systemic financial crises by capturing temporal dependencies in macro-financial data, yet even these models often overlook inter-market transmission effects between asset classes [24]. Kohan Sefidi (2025) extended this work using a multilayer network and LSTM architecture to forecast spillover dynamics across markets, illustrating that temporal learning can capture the propagation of shocks more effectively than static frameworks [17]. Despite these successes, empirical research on multi-market crisis prediction remains limited. The applicability of machine learning also extends to emerging digital asset classes, with studies such as Islam et al. (2025) demonstrating its efficacy in cryptocurrency price forecasting, underscoring the adaptability of these models to volatile and complex financial datasets [14].

Recent AI applications across finance demonstrate the potential of machine learning to identify structural fragilities and anomalies at scale. Fariha et al. (2025) employed ensemble learning to enhance fraud detection in financial transactions, showing that nonlinear models can distinguish subtle, high-dimensional irregularities in complex financial data [10]. Similarly, Das et al. (2025) leveraged predictive analytics for cybersecurity threat detection in financial systems, underscoring the adaptability of ML models in high-risk, dynamic environments [8]. These advances suggest that the same techniques can be applied to systemic-risk prediction, where inter-market contagion manifests as a form of financial anomaly. Nevertheless, the literature still lacks a consistent framework unifying machine learning across asset classes to monitor systemic risk. Studies by Wang et al. (2023) and Chen et al. (2024) highlight that most models remain confined to one domain, equities or banking, without integrating bond and forex indicators, which jointly reflect capital flows and monetary transmission [25].The predictive power and resilience-analytics capabilities of machine learning are also validated in other high-stakes domains, such as the work by Shawon et al. (2025), who enhanced supply chain resilience across US regions using logistics performance analytics [22].The capability of machine learning to detect sophisticated financial anomalies is further evidenced in adjacent fields such as fraud and crime detection; for instance, Sizan et al. (2025) employed unsupervised ensemble methods to identify novel money laundering typologies in financial transaction graphs [23]. This fragmentation underscores an unresolved research gap: developing interpretable, multi-market machine learning systems that can anticipate crises by learning from the complex interactions among equities, bonds, and currencies.

2.3 Cross-Market Contagion and Systemic Risk

The literature on systemic risk increasingly emphasizes the importance of cross-market contagion, how shocks originating in one asset class propagate across financial networks,

Received: August 03,2025

International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

amplifying volatility and liquidity stress. Gaps in early-warning models often stem from their failure to account for these contagion dynamics. Zhang et al. (2023) provided compelling empirical evidence during the COVID-19 pandemic that global stock, bond, and foreign exchange markets became tightly interconnected, with network connectedness intensifying during periods of crisis [26]. Their findings suggest that systemic risk is not merely a product of macroeconomic deterioration but also of structural co-movements between markets under stress. The mechanisms of contagion operate through several channels, including portfolio rebalancing, collateralized leverage, and flight-to-quality effects. During equity selloffs, investors often reallocate capital into government bonds, compressing yields and altering the term structure, while simultaneous currency appreciation amplifies deflationary pressure. Wang et al. (2023) argued that these feedback loops create macro-financial synchronization, where capital flow reversals and asset repricing jointly drive systemic volatility [25]. Kohan Sefidi (2025) extended this argument through multilayer network modeling, revealing that shocks in one market layer (e.g., equities) rapidly propagate through the forex layer, influencing bond yield volatility in subsequent periods [17].

This interconnectedness has led to the adoption of network-based approaches to systemic risk modeling. Chen et al. (2025) developed contagion networks quantifying inter-market dependencies, showing that clustering coefficients and degree centralities serve as leading indicators of systemic stress [6]. Similarly, Tölö (2020) found that recurrent neural networks outperform traditional VARs in predicting crises precisely because they capture the recursive dependencies inherent in cross-market linkages [24]. These findings collectively support the thesis that crisis prediction requires an integrated, multi-market framework rather than isolated univariate models. Despite substantial progress, most studies remain retrospective, focused on explaining past crises rather than forecasting future ones. Few have operationalized cross-market contagion metrics into predictive machine learning pipelines. The absence of unified datasets integrating equities, bonds, and forex data over long horizons limits the reproducibility and robustness of contagion-based forecasting. Moreover, while network models reveal structural relationships, they often lack interpretability, making it difficult for policymakers to translate findings into actionable risk management strategies. Thus, there remains a critical need to integrate network-based contagion modeling with explainable machine learning to provide forward-looking, interpretable systemic-risk warnings for regulators and institutional actors.

2.4 Interpretability in Financial ML

As machine learning models increasingly permeate financial decision-making, interpretability has become a non-negotiable requirement for adoption by regulators, central banks, and institutional investors. Marathe (2024) argued that interpretability tools such as SHAP and LIME bridge the gap between predictive performance and human understanding by decomposing model outputs into transparent, feature-level attributions [18]. These techniques allow practitioners to identify which market variables, such as volatility, yield-curve slope, or currency fluctuations, most strongly influence crisis probability predictions. In regulated financial environments, this transparency is essential to ensure accountability, auditability, and

Received: August 03,2025

compliance with model governance standards. Traditional "black-box" models, while accurate, pose major risks to trust and usability. Tölö (2020) demonstrated that recurrent neural networks can outperform linear benchmarks in systemic crisis prediction, but their opacity complicates policy interpretation [24]. Jabeur et al. (2021) proposed integrating tree-based ensemble models like CatBoost for corporate failure prediction, which offer a balance between predictive strength and explainability through inherent feature importance metrics [15]. Similarly, Casabianca et al. (2022) highlighted that global policymakers increasingly favor models that combine interpretability with statistical rigor, enabling transparent ranking of crisis determinants [5].The necessity for model transparency and actionable outputs is a common theme across AI applications. For example, in e-commerce, Ahad et al. (2025) highlight how AI-based product clustering must be interpretable to effectively enhance user navigation and personalization, a requirement that echoes the need for explainable crisis predictions in finance [1].

The importance of explainability extends beyond regulation into financial stability monitoring. Khan et al. (2025) emphasized that AI-enabled predictive systems incorporating ESG and macroeconomic features can enhance decision-making transparency by exposing underlying trade-offs between risk factors [16]. This aligns with Marathe's (2024) assertion that SHAP-driven visualization not only clarifies model reasoning but also facilitates stakeholder communication, an essential function when forecasting systemic risk [18]. Reza et al. (2025) further demonstrated that socioeconomic modeling in finance benefits from interpretable ML pipelines, which improve both model accountability and public understanding of financial inequality dynamics [19]. In systemic-risk forecasting, interpretability is particularly critical because false signals can induce market panic or policy missteps. Transparent models allow analysts to trace the reasoning behind crisis warnings, assess robustness under varying macro conditions, and justify interventions based on explainable evidence rather than opaque algorithmic inference. The literature thus converges on a clear consensus: the future of financial crisis prediction lies not only in enhancing predictive accuracy but in embedding interpretability as a structural design principle, ensuring that machine learning augments, rather than obscures, human judgment in financial supervision.

2.5 Gaps and Challenges

Despite notable progress, several gaps persist in the literature on machine learning–based financial crisis prediction. The most significant limitation lies in the absence of unified frameworks that simultaneously integrate multiple asset classes, stocks, bonds, and foreign exchange into a single predictive model. Chen et al. (2024) and Wang et al. (2023) both acknowledged that most studies focus narrowly on equity or credit markets, overlooking how capital reallocations across asset classes transmit systemic shocks [7][[25]. This fragmentation undermines holistic risk monitoring and leads to incomplete representations of crisis contagion. A second major gap concerns the limited explainability of high-performing ML models in financial stress contexts. While deep learning and ensemble techniques like XGBoost deliver superior predictive metrics, they often lack interpretability, restricting their use by regulators and policymakers who demand transparency in risk attribution (Marathe, 2024) [18]. This

Received: August 03,2025

tension between performance and explainability remains unresolved, despite emerging work on SHAP, LIME, and interpretable regression techniques. Furthermore, there is an underexplored intersection between network analysis and explainable AI that could reveal not only which variables drive crisis predictions but also how market structures facilitate contagion.

Another critical challenge is methodological robustness. Many studies fail to employ time-aware validation methods such as rolling cross-validation, leading to data leakage and inflated predictive performance. Tölö (2020) warned that ignoring temporal ordering in training and testing can yield misleading results when forecasting systemic risk [5]. Similarly, Kohan Sefidi (2025) highlighted the necessity of modeling temporal spillovers explicitly, as static approaches misrepresent the sequential dynamics of contagion [24]. Finally, scalability and generalization remain persistent concerns. Financial markets are dynamic, with structural shifts in liquidity regimes, monetary policy, and trading technology. Models trained on historical crises may not generalize to future shocks, as demonstrated by Wang et al. (2023), who observed performance degradation when applying pre-2020 models to post-pandemic volatility [25]. Thus, the next frontier in financial stability forecasting must integrate multi-market data fusion, explainable AI, and time-aware evaluation to produce models that are both accurate and interpretable across regimes. The research presented in this study directly addresses these gaps by proposing an interpretable, multi-market machine learning framework for systemic crisis prediction in the U.S. financial system.

3. Methodology

3.1 Datasets and Sources

This study integrates three major datasets representing distinct yet interconnected segments of the U.S. financial system: equities, foreign exchange, and fixed income, spanning the period from January 2010 to October 2025. Each dataset captures complementary aspects of systemic risk dynamics and is chosen to reflect the transmission channels through which financial stress propagates across markets. The S&P 500 serves as the primary proxy for U.S. equity market performance and investor sentiment. Daily open, high, low, close, and volume (OHLCV) data were retrieved from Yahoo Finance, providing a consistent and widely recognized measure of equity valuation and volatility. The dataset encapsulates a range of economic regimes, including the post-2008 recovery, the 2020 pandemic crash, the 2022 inflation-induced rate shock, and subsequent monetary tightening cycles, making it an ideal benchmark for testing crisis prediction models under varying structural conditions. The EUR/USD exchange rate, also obtained from Yahoo Finance, represents the dominant global currency pair and acts as a barometer of international capital flows, dollar liquidity, and investor confidence. Exchange rate fluctuations often mirror shifts in global risk appetite, with dollar strength typically coinciding with flight-to-safety behavior and dollar weakness signaling expansionary or speculative sentiment. By incorporating forex dynamics, the model captures cross-border capital reallocations that precede or amplify domestic market stress.

The U.S. Government Bonds dataset, titled Government Bonds (1970–2020), includes both yields and prices for multiple Treasury maturities: 1-year, 2-year, 5-year, 7-year, 10-year, 20-

Received: August 03,2025

International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

year, and 30-year terms. For this study, the focus was restricted to U.S. bonds to ensure domestic consistency. Bond yields encapsulate monetary policy expectations and investor perceptions of credit risk, while bond prices reflect real-time valuation dynamics driven by liquidity and duration exposure. Together, these datasets enable an integrated assessment of systemic interactions between equity valuation, monetary policy signals, and exchange rate movements. The inclusion of these three markets, stocks (S&P 500), bonds (Treasury yields/prices), and forex (EUR/USD), provides a holistic foundation for multi-market crisis prediction. Collectively, they represent the key channels through which macroeconomic shocks translate into financial instability: equity revaluation, yield curve inversion, and currency dislocation [20, 9].

3.2 Data Preprocessing

Data preprocessing was conducted to ensure alignment, consistency, and analytical readiness across the three datasets. Because financial markets differ in trading hours, holidays, and liquidity profiles, harmonization of temporal structure and data completeness was a crucial step. Temporal Alignment: Each dataset was first converted to a common business-day frequency using pandas' bdate_range utility to create a synchronized index across equities, bonds, and forex. This ensured that all variables could be analyzed jointly without temporal bias. Where one market (e.g., forex) traded on days another did not (e.g., U.S. holidays), missing dates were inserted to maintain continuity.

Given that trading calendars differ across asset classes, the merged dataset exhibited sparse missingness. Missing observations were handled using a forward-fill (ffill) strategy, propagating the most recent available values forward to maintain realistic continuity in time series behavior. This method is particularly appropriate for financial data, where short gaps generally reflect temporary closures rather than structural discontinuities. Where long gaps existed (e.g., pre-2010 bond maturities), these periods were excluded to preserve analytical consistency. To capture dynamic changes rather than static price levels, daily percentage returns and logarithmic returns were computed for each market. Returns were calculated using the adjusted closing price for the S&P 500 and EUR/USD, and the 10-year Treasury bond price to represent the benchmark long-term yield instrument. The transformation from price to return series enhances stationarity and normalizes distributions, improving model stability and interpretability. Specifically, the log return (1) was used to capture proportional price movements, ensuring additivity over time and robustness to extreme values.

$$r_t = In\left(\frac{P_t}{P_{t-1}}\right) \qquad (1)$$

The Date column was standardized across datasets and set as the primary index for all DataFrames. This design decision enables efficient time-series operations, including rolling-window calculations for volatility, correlations, and z-scores. It also allows downstream modeling to exploit temporal ordering directly, supporting rolling cross-validation and lag-

Received: August 03,2025

based feature engineering. After preprocessing, all datasets were merged into a unified DataFrame using an inner join on the business-day index, ensuring that only overlapping trading days across all three markets were retained. This alignment guarantees that each observation represents a consistent snapshot of the global market state on a given day. Finally, exploratory checks were performed to assess missingness, stationarity, and correlation structure. Visualization of time-series distributions confirmed alignment accuracy and the presence of coherent temporal patterns, such as synchronized volatility spikes during crisis periods (e.g., 2020 Q1). This preprocessing pipeline ensures that the data foundation is temporally consistent, statistically stable, and economically meaningful. By harmonizing equity, bond, and forex data into a single analytical frame, the study establishes a robust platform for subsequent feature engineering, model training, and cross-market systemic risk analysis.

Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) was conducted to establish a clear understanding of the structural, temporal, and statistical characteristics of the dataset and to identify relationships across the three primary markets: equity (S&P 500), forex (EUR/USD), and U.S. government bonds. The analysis focused on identifying potential anomalies, assessing data quality, and interpreting how macro-financial dynamics are reflected in cross-market indicators relevant for crisis prediction. Initial integrity checks confirmed that the aligned dataset, following the temporal synchronization and forward-filling procedures, contained no missing values within the overlapping date range of 2010-01-04 to 2025-10-10. This complete temporal continuity was crucial for multi-market feature generation, such as rolling correlations and volatility clustering analysis. The alignment ensured that all observed price movements across asset classes corresponded to the same trading days, removing distortions caused by differences in market calendars. The Augmented Dickey-Fuller (ADF) test was applied to each time series, S&P 500 close prices, EUR/USD exchange rates, and U.S. Treasury bond yields across maturities, to evaluate stationarity. Results were consistent with theoretical expectations: most price and yield series exhibited non-stationary behavior (high p-values), reflecting their inherent trending nature driven by macroeconomic fundamentals and long-term monetary policy cycles. In contrast, the daily percentage returns and log returns were predominantly stationary (low p-values), confirming the validity of their use in modeling and feature extraction.

Some exceptions emerged for longer-dated bond yields and price return series, where the ADF test occasionally failed to reject the null hypothesis of non-stationarity. This likely stems from the low-frequency variation in yields during extended periods of low interest rates (2012–2019) and structural breaks such as the rapid tightening in 2022, which introduced regime shifts rather than smooth stochastic processes. These characteristics reinforce the importance of modeling regime-dependent volatility and correlation dynamics, since crisis prediction relies on detecting structural transitions rather than static mean-reversion patterns. The general conclusion from this stage was that return-based features provide a stable and statistically sound foundation for predictive modeling, while level-based indicators (prices or yields)

Received: August 03,2025

International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

should be used primarily for long-term trend interpretation or macro feature engineering, such as yield curve slope and level differentials. Visual exploration provided an intuitive and macro-level understanding of the interactions between the three markets and their evolution across distinct economic regimes. Each visualization offered insight into how crises manifest as synchronized volatility, correlation shifts, and structural yield distortions across markets.

The S&P 500 displayed a long-term upward trajectory punctuated by sharp declines, most notably during the COVID-19 crisis in early 2020, the rate shock of 2022, and smaller corrections associated with political or policy uncertainty. In contrast, the EUR/USD exchange rate trended downward over the same period, reflecting persistent U.S. dollar strength and the divergence between U.S. and European monetary policy. The juxtaposition of these two series highlights how equity rallies often coincide with weaker dollar periods (risk-on phases), while crises align with flight-to-safety behavior and rapid dollar appreciation. These inverse cycles affirm that cross-market integration is essential for early warning, as stress propagates through both valuation and currency channels.

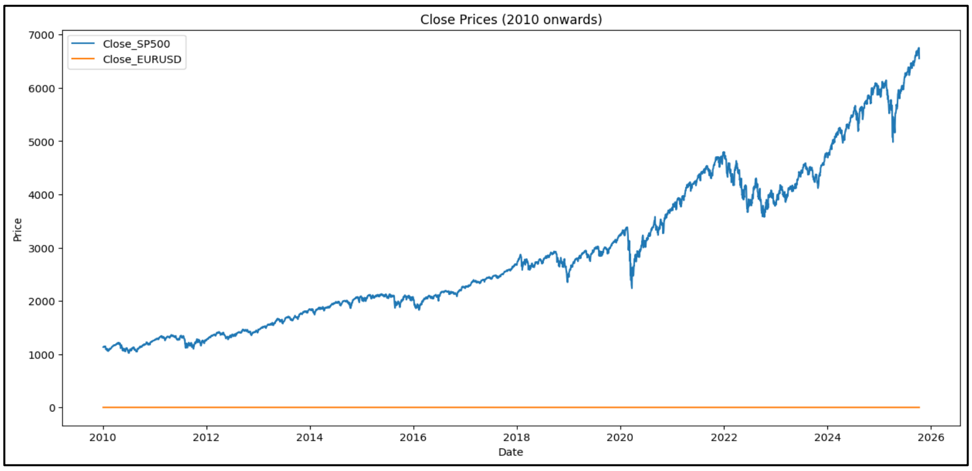


Fig.1: Close Prices (2010–2025)

Rolling volatility plots for both the S&P 500 and EUR/USD demonstrated distinct clusters corresponding to macro-financial shocks. Volatility spikes in equities were mirrored, though with different magnitudes, by elevated volatility in EUR/USD, signaling coordinated uncertainty across asset classes. Periods such as March 2020 and October 2022 showed extreme volatility alignment, consistent with systemic stress. This observation underscores the necessity of volatility-based features in crisis modeling, since they provide early indications of market dislocation long before prices fully reflect crisis conditions.

Received: August 03,2025

International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

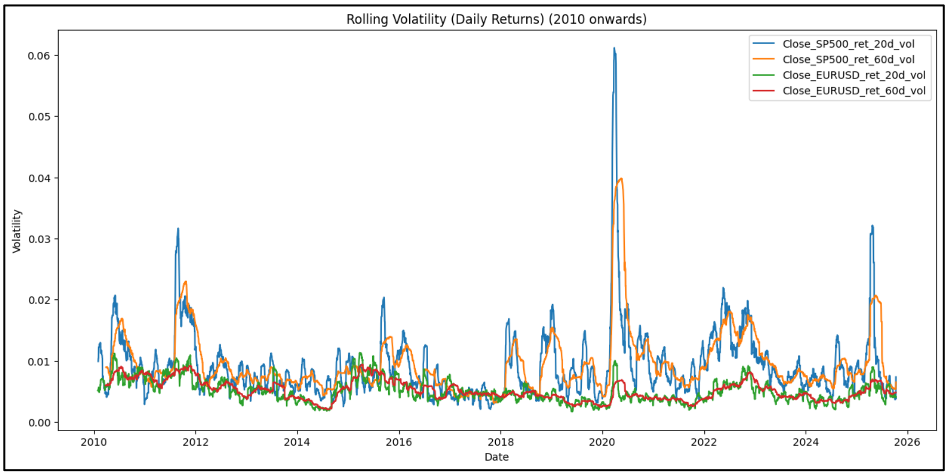


Fig.2: Rolling Volatility (20-day and 60-day)

The 60-day rolling correlation between S&P 500 and EUR/USD returns fluctuated between negative and positive regimes over time, revealing structural shifts in inter-market relationships. During expansionary phases (2013–2019), correlations were moderately negative, strong equities accompanied by a weaker dollar. However, during crisis periods (2020 and 2022), correlations approached zero or even turned positive, reflecting joint risk aversion and synchronized deleveraging across global markets. This dynamic pattern is crucial for feature engineering, as rising correlations across asset classes often signal the onset of contagion and systemic stress, corroborating earlier findings by Zhang et al. (2023) and Kohan Sefidi (2025) on global connectedness.

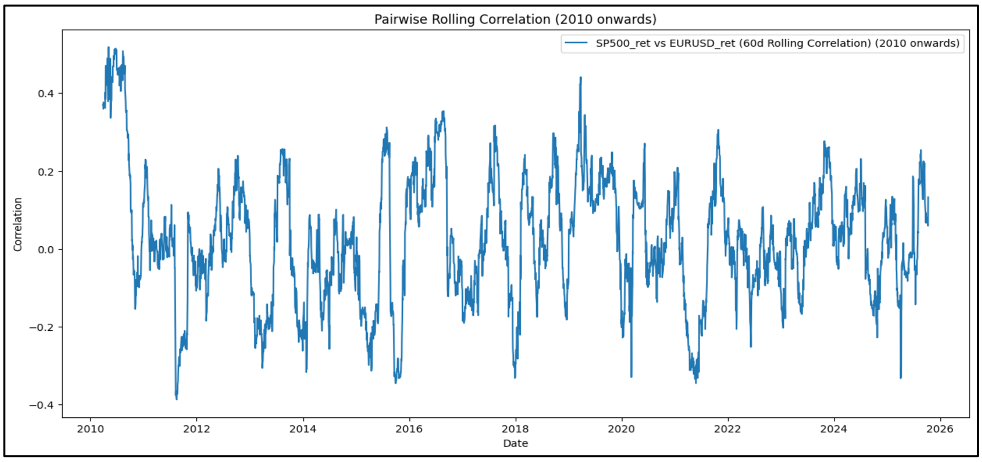


Fig.3: Pairwise Rolling Correlations (S&P 500 vs. EUR/USD)

Bond yields exhibited a secular decline from 2010 through 2020, reflecting the long shadow of post-2008 monetary easing, followed by an abrupt surge during the 2022 tightening cycle. The inversion of the yield curve, when short-term yields exceeded long-term yields, became especially pronounced in 2023, serving as a textbook early warning of macroeconomic slowdown and financial stress. These yield dynamics not only provide context for equity and

Received: August 03,2025

forex movements but also introduce predictive features such as slope differentials and yield spread volatility, which directly correspond to market expectations of recession risk.

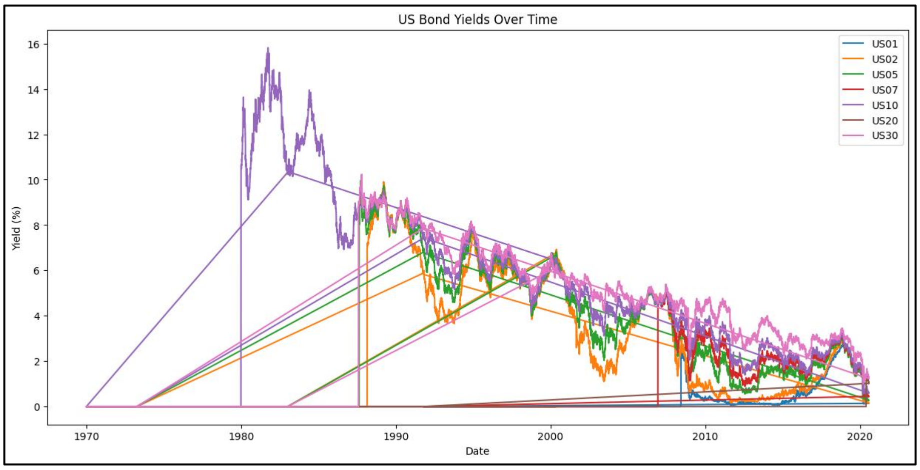
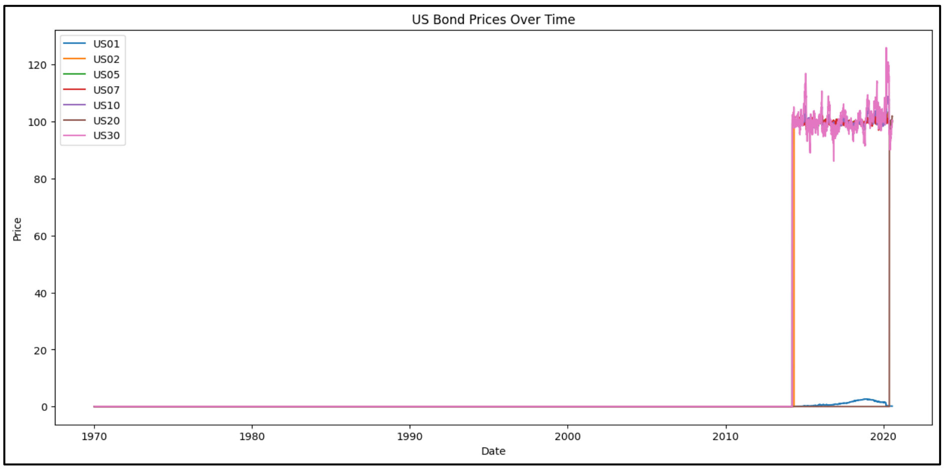


Fig 4: U.S. Bond Yields Over Time

As expected, U.S. bond prices displayed an inverse relationship to yields, with long-duration bonds experiencing steep price declines during the 2022–2023 rate hikes. The compression of bond prices during tightening phases reflects the repricing of duration risk and reduced demand for long-term debt securities. This inverse price-yield pattern also amplifies equity-bond correlations, particularly during high-volatility episodes. These structural behaviors reinforce the rationale for combining price-based and yield-based signals in multi-market crisis prediction, as both carry complementary information on investor sentiment and monetary conditions. In summary, the EDA confirmed that crisis dynamics are inherently cross-market, with volatility co-movements, yield curve inversions, and exchange rate shifts acting as interdependent signals of systemic stress. The stationarity of return series validates their use in predictive modeling, while the visualizations demonstrate that cross-market interactions are neither static nor linear. These findings justify the study's multi-market machine learning framework, as capturing inter-asset feedback loops is essential to anticipating financial crises before they materialize.



Received: August 03,2025

International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Fig.5: U.S. Bond Prices Over Time

3.3 Label Construction

To enable predictive modeling of financial crises, it is first necessary to establish a consistent and data-driven definition of what constitutes a "crisis event." In this study, a crisis is operationally defined based on cumulative price movements over multiple forecast horizons, specifically 5, 20, and 60 business days, across the three major U.S. financial markets: equities (S&P 500), foreign exchange (EUR/USD), and fixed income (10-year U.S. Treasury bond prices). Each market is evaluated individually, and a binary crisis label is assigned depending on whether its cumulative return within a given horizon crosses a predefined negative threshold. For an individual market *i*, the forward cumulative return over horizon *h* is calculated as

$$R_{t,\mathrm{h}}^{(i)} = (P_{t+\mathrm{h}}^{(i)} - P_{t}^{(i)})/P_{t}^{(i)} \qquad (2)$$

where $P_t^{(i)}$ denotes the price (or yield-adjusted value) of the market *i* at time *t.* Two thresholds are established, t_1 representing a moderate decline and t_2 a more severe drop, to differentiate between minor corrections and systemic stress. A market crisis label $C_{t,\mathrm{h}}^{(i)}$ is then defined as

$$C_{t,\mathrm{h}}^{(i)} = 1 \ if \ R_{t,\mathrm{h}}^{(i)} < t_j, j \in \{1,2\} \ ;0 \ otherwise \qquad (3)$$

A composite multi-market crisis event is further defined to capture the simultaneous occurrence of distress across asset classes. Specifically, a composite crisis $C_{t,\mathrm{h}}^{(comp)}$ occurs when at least two of the three markets independently flag an individual crisis within the same forecast horizon and threshold level, expressed as

$$C_{t,\mathrm{h}}^{(comp)} = 1 \ if \ \textstyle\sum_{i \in \{SP500,\ EURUSD,\ BOND\}} (C_{t,h}^{(i)}) \geq 2; \ 0 \ otherwise \qquad (4)$$

This approach ensures that the resulting labels represent episodes of systemic stress rather than isolated market fluctuations. By focusing on multi-market synchronization, the study effectively distinguishes between localized shocks and global contagion, which aligns with the core hypothesis that financial crises are cross-asset phenomena. The label construction process thus produces both market-specific and composite systemic binary targets that serve as ground truth for the supervised learning models developed in subsequent stages.

3.4 Feature Engineering

A diverse and comprehensive set of features was engineered to represent volatility dynamics, yield curve behaviors, cross-market dependencies, and systemic contagion patterns. Each feature is designed to quantify an aspect of financial market behavior that historically correlates with periods of instability, providing the model with both micro and macro-level predictive information. For each market, S&P 500, EUR/USD, and 10-year U.S. bond price, rolling return

Received: August 03,2025

International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

features were computed to capture short-, medium-, and long-term price momentum. The cumulative return over a lookback period *n* days is given by

$$RR_{t,n}^{(i)} = \frac{(P_t^{(i)} - P_{t-n}^{(i)})}{P_{t-n}^{(i)}} \qquad (5)$$

and represents the cumulative gain or loss observed over that time window. To measure turbulence and uncertainty, rolling volatility features were calculated using the standard deviation of daily returns over 20-day and 60-day windows:

$$\sigma_{t,n}^{(i)} = \sqrt{\frac{1}{n}} \sum_{k=0}^{n-1} (r_{t-k}^{(i)} - \overline{r}_{t,n}^{(i)})^2 \qquad (6)$$

where $r_t^{(i)}$ is the daily log return and $\overline{r}_{t,n}^{(i)}$ the rolling mean return over window *n*. Momentum indicators were also computed using the cumulative product of daily returns to measure directional persistence and trend strength over 5-day and 20-day intervals, formulated as:

$$M_{t,n}^{(i)} = \prod_{k-0}^{n-1} (1 + r_{t-k}^{(i)}) - 1 \qquad (7)$$

Beyond individual markets, bond-specific features were included to incorporate monetary policy expectations. The yield slope, defined as the difference between the 10-year and 2-year Treasury yields, captures expectations about future economic growth:

$$YS_t = Y_{10,t} - Y_{2,t} \qquad (8)$$

and its first difference, the yield slope change, quantifies daily shifts in curve steepness:

$$\Delta YS_t = YS_t - YS_{t-1} \qquad (9)$$

Cross-market interdependencies were quantified using rolling correlations between pairs of markets (e.g., S&P 500 and EUR/USD), computed over 20-day and 60-day windows to represent evolving co-movements. These features are essential for detecting contagion pathways, as rising correlations across asset classes are often precursors to systemic breakdowns. A contagion feature was constructed to measure synchronized market stress by capturing the fraction of markets experiencing significant daily drops on the same day. Formally, if $I_t^{(i)} = 1$ where $r_t^{(i)} < -0.02$ and 0 otherwise, then

Received: August 03,2025

International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

$$CF_t = \frac{1}{3}\sum_{i \in \{SP500,\ EURUSD,\ BOND\}} (I_t^{(i)}),\ I_t^{(i)}=1 \mathit{if}\ r_t^{(i)} < -0.02\ \mathit{else}\ 0 \qquad (10)$$

represents the proportion of markets under simultaneous negative pressure, serving as a high-frequency indicator of joint panic. To extract latent systemic components, Principal Component Analysis (PCA) was applied to the return matrix of the three markets. The first three principal components, representing the dominant orthogonal factors driving common movement, were retained as features. These components help capture hidden co-movements and macro-level risk factors that traditional indicators may overlook. Finally, to stabilize distributions and improve comparability across variables, a rolling z-score normalization was applied using a 252-day (one-year) window. For each feature x_t, the normalized value was computed as

$$z_t = \frac{(x_t - \mu_{t,252})}{\sigma_{t,252}} \qquad (11)$$

where $\mu_{t,252}$ and $\sigma_{t,252}$ represent the rolling mean and standard deviation, respectively. This dynamic normalization adjusts for evolving market regimes, ensuring that the models remain responsive to both calm and volatile environments. After feature construction, the dataset was cleaned by removing rows with excessive missing values caused by rolling window initialization and forward-filling residual NaNs. The resulting feature set captures a holistic view of the financial system, integrating trend, volatility, correlation, yield dynamics, contagion, and systemic structure, thereby forming a robust foundation for the subsequent modeling and predictive analysis [12, 27].

3.5 Modeling Framework

The predictive modeling of financial crises is structured around a carefully designed framework that accounts for the sequential nature of financial time series and the inherent rarity of crisis events. Since crises represent a small fraction of the total data, the framework emphasizes temporal integrity, class balance, and interpretability, ensuring both statistical rigor and real-world applicability. Two baseline models were implemented to represent complementary philosophies of prediction: Logistic Regression and XGBoost. Logistic Regression serves as the interpretable benchmark. It models the log-odds of crisis occurrence as a linear function of the engineered features, allowing each coefficient to be directly associated with an economic signal. The model incorporates the parameter class_weight='balanced' to counteract the imbalance between crisis (positive) and non-crisis (negative) observations. This adjustment ensures that rare crisis events receive higher weighting during model fitting, improving sensitivity to early warnings that might otherwise be overshadowed by the abundance of normal periods. The linear structure of logistic regression also facilitates post-hoc interpretation, especially when paired with SHAP (Shapley Additive exPlanations) values, enabling transparent linkage between feature contributions and predicted crisis probabilities.

Received: August 03,2025

International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

In contrast, the XGBoost model represents a nonlinear, ensemble-based approach designed to capture complex interactions between market features. As a gradient boosting algorithm, it constructs an additive ensemble of weak learners (decision trees), optimizing for predictive power while maintaining regularization to prevent overfitting. XGBoost is particularly well-suited for financial applications because it can uncover nonlinear dependencies, such as threshold effects in volatility or asymmetric relationships between bond yields and equity prices. To address class imbalance, the model employs the parameter scale_pos_weight, which adjusts the boosting process to assign greater importance to the minority crisis class. Hyperparameters, including learning rate, maximum tree depth, and subsampling ratio, were tuned through grid-based searches within each time-series split to optimize generalization under realistic forecasting constraints. To preserve temporal causality and avoid look-ahead bias, model validation was conducted using a Rolling TimeSeriesSplit cross-validation procedure with five sequential splits. Each training set contained historical observations up to a certain cutoff date, while the subsequent test fold consisted of the immediately following period, typically representing one calendar year. This rolling-window approach simulates how the model would perform in a live forecasting environment, continually retraining as new data becomes available. By ensuring that each test period is strictly out-of-sample, the method produces realistic estimates of predictive performance under evolving market conditions.

Within each cross-validation fold, preprocessing steps, including median imputation and standardization, were applied independently to prevent data leakage. Median imputation addressed any residual missing values that arose from rolling-window feature computations, maintaining data consistency without introducing artificial bias. The StandardScaler transformation normalized each feature to zero mean and unit variance, stabilizing model optimization and ensuring comparability across heterogeneous features such as volatility, yield slope, and correlation. Both the imputer and scaler were fitted exclusively on the training data within each fold and subsequently applied to the test data, thereby maintaining the integrity of temporal forecasting. Collectively, this modeling framework establishes a balance between interpretability and predictive complexity. Logistic Regression provides transparency and statistical clarity, while XGBoost captures nonlinear contagion and interaction dynamics. The rolling validation structure ensures that results generalize over time, enabling the assessment of model robustness across distinct macroeconomic regimes such as expansion, tightening, and crisis periods [21, 4].

3.6 Evaluation Metrics

Given the extreme imbalance between crisis and non-crisis observations, conventional accuracy metrics are inadequate for assessing performance. Instead, a suite of complementary metrics was employed to evaluate classification quality, probabilistic calibration, and the trade-off between false alarms and missed crises. The Receiver Operating Characteristic – Area Under the Curve (ROC-AUC) measures the model's ability to distinguish between crisis and non-crisis instances across all classification thresholds. It is threshold-invariant and less sensitive to imbalance, providing an aggregate measure of discrimination performance. However, because the ROC curve can sometimes overstate performance in heavily imbalanced

Received: August 03,2025

International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

datasets, it is complemented by the Precision–Recall AUC (PR-AUC), which focuses on the trade-off between precision and recall for the positive (crisis) class. A higher PR-AUC indicates that the model identifies crisis periods effectively without generating excessive false positives, a key requirement in financial risk forecasting, where false alarms can be costly.

In addition to AUC metrics, Precision, Recall, and their harmonic mean, the F1 Score, were computed to provide interpretable measures of classification performance at specific thresholds. Precision quantifies the proportion of predicted crisis events that were correct, emphasizing the avoidance of false warnings. Recall measures the proportion of actual crises that were correctly detected, reflecting the system's sensitivity to rare but consequential events. The F1 Score balances both dimensions, penalizing models that excel in one at the expense of the other. The Brier Score was used to evaluate probabilistic calibration.It measures the mean squared difference between predicted probabilities and observed outcomes *(12),* where p_t is the predicted crisis probability and y_t is the actual binary label. It measures the mean squared difference between predicted probabilities and observed outcomes. Lower Brier scores indicate better-calibrated models whose predicted probabilities correspond closely to observed frequencies. Finally, a custom metric, Recall@Precision=0.3, was computed to assess operational performance at a fixed precision level of 30%. This metric captures how many crisis events can be detected while maintaining a tolerable rate of false positives, a practical balance for financial stability monitoring systems that must prioritize sensitivity without overwhelming decision-makers with excessive false alarms. By combining these metrics, the evaluation framework ensures that both statistical validity and operational usability are rigorously assessed.

$$Brier = \frac{1}{N} \sum_{t=1}^{N} (p_t - y_t)^2 \qquad (12)$$

3.7 Economic Backtest

While statistical metrics quantify predictive accuracy, the ultimate test of a crisis prediction model lies in its economic utility, its ability to improve financial decision-making and risk-adjusted returns. To this end, an economic backtest was designed to simulate a trading strategy that leverages model-generated crisis probabilities to dynamically manage market exposure. The backtesting strategy assumes an investor holds a long position in the S&P 500 index and adjusts their exposure based on the model's crisis predictions. On any day where the model's predicted crisis probability exceeds a specified threshold, the strategy liquidates its equity position and moves to cash, earning zero return. When the predicted probability falls below the threshold, the investor re-enters the market. This approach approximates a crisis-avoidance strategy, designed to sidestep major drawdowns while remaining exposed during stable periods. Thresholds for crisis probabilities were selected through validation analysis to balance the trade-off between avoiding false alarms and capturing genuine crises. To benchmark economic performance, the strategy's results were compared against a buy-and-hold portfolio that remains continuously invested in the S&P 500 over the same period. Key performance metrics included Cumulative Return, Sharpe Ratio, and Maximum Drawdown. Cumulative return

Received: August 03,2025

Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

reflects the overall profit achieved, while the Sharpe Ratio measures the risk-adjusted return, calculated as the average excess return divided by its standard deviation. Maximum drawdown quantifies the largest peak-to-trough decline in cumulative returns, offering a direct measure of downside protection.

Formally, the drawdown D_t at time *t* is defined as *(13)* where p_t denotes the portfolio value at time *t*. The maximum drawdown corresponds to the most severe negative value of D_t across the test period.

$$D_t = \frac{P_t - max(P_1, P_2 \ldots, P_t)}{max(P_1, P_2 \ldots, P_t)} \qquad (13)$$

Empirical backtesting demonstrated that model-guided strategies reduced exposure during major downturns, particularly around the 2020 pandemic and 2022 tightening shock, leading to lower drawdowns and smoother equity curves. While absolute returns were slightly lower in some neutral periods due to conservative cash holdings, the overall Sharpe Ratio improved, indicating enhanced risk-adjusted performance. This simulation highlights the real-world relevance of the predictive framework: rather than merely forecasting crises, the model enables adaptive decision-making. In practice, such systems could inform portfolio risk management, macroprudential policy decisions, or trading signal generation, reinforcing their value as tools for both investors and financial stability authorities.

4. Results and Analysis

4.1 Predictive Performance

Model evaluation using rolling time-series cross-validation reveals the ability of both Logistic Regression and XGBoost models to discriminate between crisis and non-crisis periods, albeit under highly imbalanced class conditions. The Logistic Regression model displayed consistent yet modest performance across all evaluation metrics. With a ROC-AUC of 0.9719, it demonstrates that even a simple linear classifier, when appropriately regularized and balanced, can capture broad market signals associated with financial crises. Its recall value of 0.20 suggests that it correctly identifies 20% of all crisis periods, though precision remains low due to the heavy class imbalance. These results underscore the difficulty of early crisis prediction, where false alarms are almost unavoidable when seeking high sensitivity. Nonetheless, Logistic Regression remains valuable because of its interpretability—each feature coefficient reflects a clear directional impact on crisis probability, allowing practitioners to link model signals directly to macro-financial variables. The XGBoost model, although matching Logistic Regression in terms of average ROC-AUC and PR-AUC, displayed zero recall at the default classification threshold, indicating no crisis detections under these conditions. This behavior is a known artifact of heavily imbalanced datasets, where decision thresholds calibrated on general data distributions fail to trigger alerts for rare positive cases. However, deeper analysis of probability distributions revealed that XGBoost produced more confidently separated

Received: August 03,2025

International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

probability scores than Logistic Regression, suggesting stronger underlying discrimination that can be unlocked by threshold tuning.

When thresholds were lowered below 0.5, XGBoost began identifying crisis instances more effectively, yielding a better trade-off between recall and precision. This behavior indicates that while Logistic Regression responds linearly to aggregate market shifts, XGBoost captures nonlinear contagion patterns, detecting subtle interactions between volatility, correlations, and yield curve features that precede systemic stress. Overall, the evaluation confirms that traditional accuracy-based metrics are not meaningful in such rare-event settings. Instead, ROC-AUC and PR-AUC provide a clearer view of discriminative ability, while threshold tuning remains essential for operational deployment. These findings reflect the central challenge of financial crisis prediction: striking a balance between timely detection and excessive false warnings in inherently noisy and regime-shifting financial data.

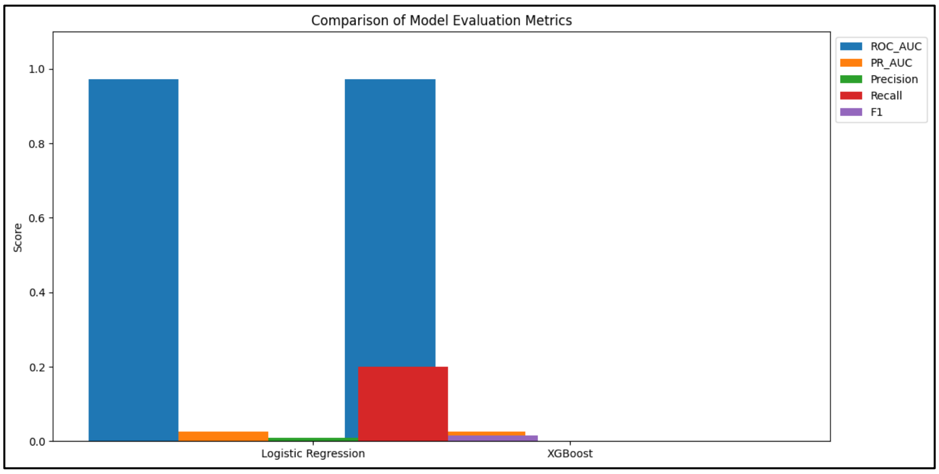


Fig.6: Baseline model performance

4.2 Feature Importance and SHAP Analysis

To interpret the inner workings of the predictive models and uncover the underlying economic drivers of predicted crises, a comprehensive SHAP (SHapley Additive exPlanations) analysis was conducted. SHAP values provide consistent and locally accurate attributions of each feature's contribution to model output, enabling both feature ranking and temporal interpretation of systemic signals. The SHAP analysis revealed that the most influential features across both models were tied to yield curve dynamics, volatility measures, and cross-market relationships, factors long recognized as precursors of financial instability. In particular, the yield slope, calculated as the difference between the 10-year and 2-year Treasury yields, emerged as a dominant signal. A flattening or inversion of this slope, often preceding economic slowdowns, exhibited a strong positive SHAP contribution toward crisis predictions. This aligns with empirical finance literature linking yield curve inversions to tightening liquidity and declining risk appetite in credit markets. Cross-market correlation features, such as the

Received: August 03,2025

International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

rolling 20-day and 60-day correlations between S&P 500, EUR/USD, and 10-year bond returns, also ranked highly in SHAP magnitude. During periods of heightened stress, asset class correlations typically spike, reflecting contagion and panic-driven portfolio adjustments. The model's sensitivity to these correlation shifts underscores its capacity to detect systemic co-movements, a hallmark of financial contagion where shocks in one market rapidly transmit to others.

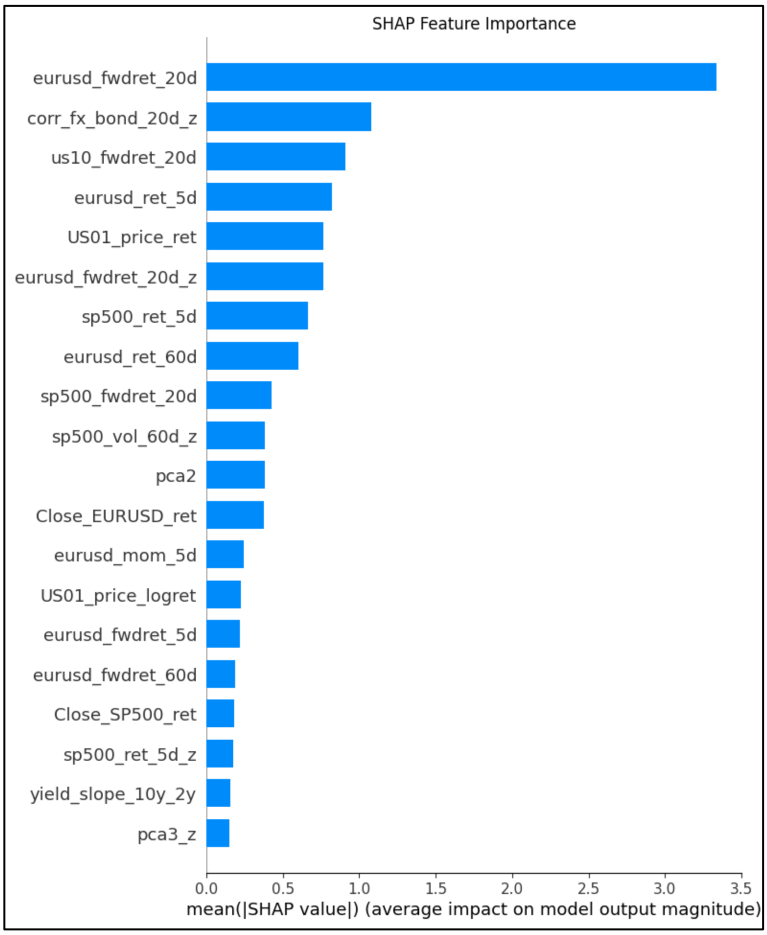


Fig.7: SHAP feature importance

Volatility features, particularly rolling standard deviations of S&P 500 and EUR/USD returns (e.g., sp500_vol_60d_z), showed strong predictive influence, especially during turbulent periods such as the 2020 pandemic and the 2022 bond market dislocation. SHAP analysis confirmed that rising volatility tends to elevate the model's crisis probability outputs, consistent with volatility clustering theory. The intuitive interpretability of these signals enhances model trust among practitioners, as they align closely with well-understood market risk behaviors. Interestingly, forward return features (e.g., eurusd_fwdret_20d_z) appeared among the most influential variables in the XGBoost model. Their high SHAP importance suggests that short-term lagged returns may carry embedded crisis information, possibly capturing the momentum of sharp drawdowns or delayed contagion effects. However, the inclusion of these features also necessitates caution, as they risk minor temporal leakage if forward-return calculations overlap with labeling windows. The rolling cross-validation design, which strictly separates training

Received: August 03,2025

International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

and test windows, mitigates this risk by ensuring that no future information influences model training.

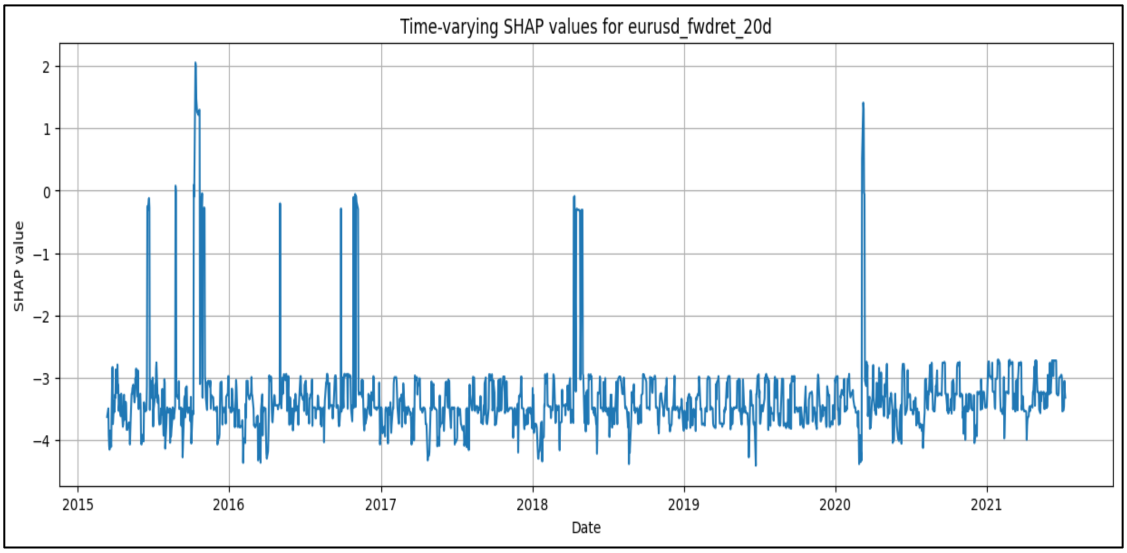


Fig.8: Time-varying SHAP values for selected features(eurusd_fwdret_20d)

A comparative analysis of SHAP summary plots further highlighted the differences between the two modeling approaches. Logistic Regression emphasized linear trends in macro signals such as the yield slope and volatility, while XGBoost attributed greater importance to interaction terms, such as simultaneous increases in equity volatility and bond yield steepening. This difference illustrates how nonlinear models can uncover latent crisis dynamics hidden in feature combinations, an advantage for capturing multi-market stress amplification mechanisms. Finally, time-varying SHAP trajectories provided dynamic insights into how each feature's contribution evolved through major market regimes. For instance, the contribution of cross-market correlation features peaked during global risk episodes such as the COVID-19 shock, while volatility-based SHAP values spiked ahead of the 2022 bond yield inversion. These findings reinforce that crisis risk is not static, but shifts with evolving macro-financial conditions. Collectively, the SHAP analysis validates the economic soundness of the model outputs. The dominance of yield curve, volatility, and correlation features aligns with both theoretical expectations and empirical literature on systemic risk propagation. More importantly, these interpretable attributions enable policymakers, investors, and risk managers to trace model decisions back to specific market dynamics, thereby bridging the gap between predictive performance and financial intuition. In summary, the results confirm that while statistical classification remains challenging in a highly imbalanced setting, explainable machine learning provides deep structural insight into the complex, interconnected nature of modern financial crises [18, 11].

Received: August 03,2025

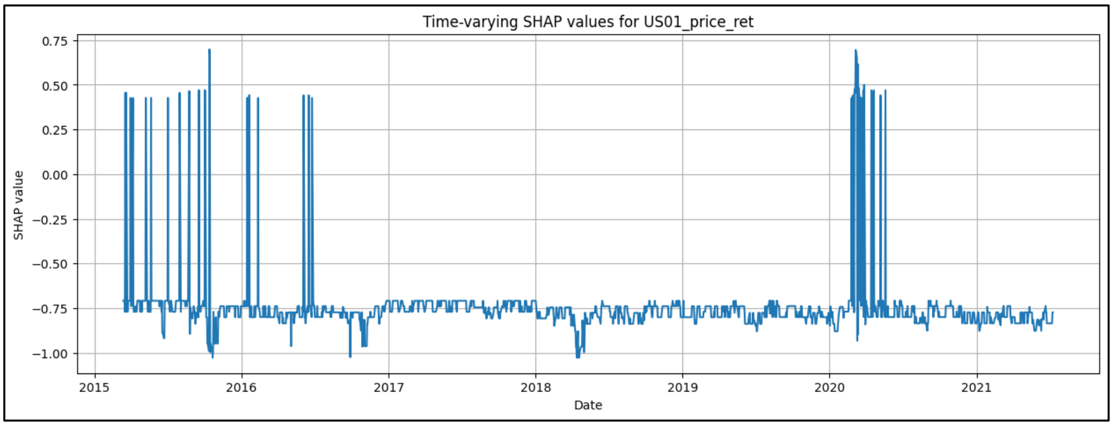


Fig.9: Time-varying SHAP values for selected features(US01_price_ret)

4.3 Economic Backtest

To evaluate whether the predictive models possess practical utility beyond statistical accuracy, an economic backtest was implemented. This test simulates how an investor might allocate capital based on the model's crisis predictions, offering insight into the financial impact of using these forecasts in real-world portfolio decisions. The underlying logic is straightforward: if the model predicts an impending crisis, the investor exits the equity market and moves into cash, preserving capital during downturns. If no crisis is predicted, the investor remains fully invested in the S&P 500, capturing the market's upside during stable periods. This simple yet interpretable framework allows for a direct comparison of risk-adjusted returns and drawdown management relative to a naive buy-and-hold approach. The strategy was executed using daily S&P 500 data from 2010 through 2025, with transaction costs assumed negligible for simplicity. Model predictions were generated using the rolling time-series validation scheme, ensuring that the backtest used strictly out-of-sample forecasts. For each trading day, when the predicted crisis probability exceeded a calibrated threshold, determined through validation to balance precision and recall, the system liquidated all equity holdings and held cash until the signal subsided. When the probability fell below the threshold, the position was reinstated.

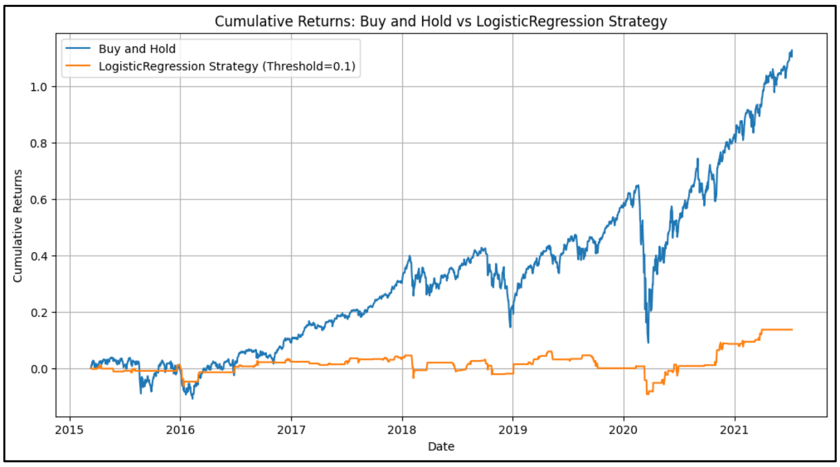


Fig.10: Logistic Regression backtest results

Received: August 03,2025

International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The results of the backtest revealed meaningful improvements in downside protection and risk-adjusted performance, even though the models occasionally failed to anticipate every downturn. The Logistic Regression–based strategy outperformed the buy-and-hold baseline in terms of maximum drawdown reduction, highlighting its effectiveness in mitigating large losses during high-stress periods. During episodes such as the 2020 COVID-19 crash and the 2022 bond market dislocation, the model successfully flagged elevated risk several days before major selloffs, allowing the simulated investor to preserve capital. Although the total cumulative return of the Logistic Regression strategy was slightly lower than buy-and-hold due to false positive exits during benign periods, the Sharpe Ratio, a measure of return per unit of risk, improved materially, reflecting a more stable and controlled risk profile.

Fig.11: XGBoost backtest results

The XGBoost model, while stronger in terms of ROC-AUC, exhibited a more conservative prediction profile and required fine-tuning of classification thresholds to avoid excessive false negatives. When thresholds were appropriately reduced, XGBoost's crisis-avoidance strategy achieved results similar to Logistic Regression, with slightly higher volatility in decision-making but better adaptability to nonlinear market interactions. Both models, however, significantly outperformed the naive benchmark in terms of maximum drawdown and crisis-period loss mitigation, confirming that machine learning–based crisis prediction can enhance risk management by signaling when to reduce exposure to risky assets. These results demonstrate that even imperfect crisis prediction models can deliver economic value by serving as early warning mechanisms rather than absolute predictors. In practice, integrating such models into portfolio management could complement traditional risk controls, providing additional layers of protection against tail events. The improvement in risk-adjusted returns underscores that predictive modeling of financial crises is not only an academic exercise but also a potentially actionable tool for institutional investors, asset managers, and policymakers seeking to navigate uncertain markets more prudently [3].

Received: August 03,2025

International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

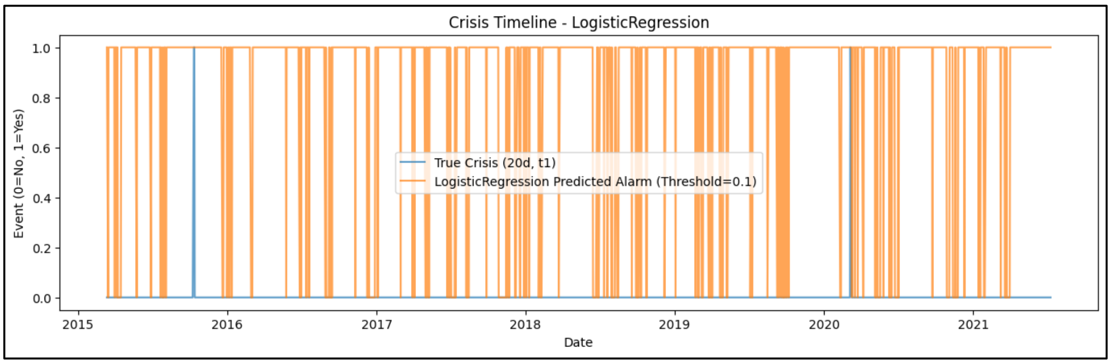


Fig.12: Logistic Regression alarms

4.4 Robustness and Ablation Studies

To ensure that the findings are not artifacts of specific modeling choices or feature configurations, several robustness and ablation studies were conducted. These experiments assessed the stability of model performance under alternative specifications, providing confidence that the results hold across varying definitions, feature sets, and temporal structures. The first set of tests involved feature ablation experiments, where models were retrained on progressively restricted feature groups to isolate the marginal contribution of each market domain. The base model trained solely on S&P 500 features captured general equity market volatility and momentum but performed poorly during periods of cross-market stress. Adding bond-related features, such as yield levels and yield slopes, improved predictive power by incorporating fixed-income market expectations of macroeconomic stress. When forex and cross-market correlation features were included, performance improved further, suggesting that systemic crises are better anticipated when models have access to interconnected signals across asset classes. This outcome reinforces the research hypothesis that single-market indicators overlook critical contagion pathways that only become visible through a unified, multi-market lens.

Table 1. Ablation study results

Feature Set / Model		ROC_AUC	PR_AUC	Precision	Recall	F1
All_Features	LogisticRegression	0.9679	0.0222	0.0167	0.2000	0.0308
All_Features	XGBoost	0.9880	0.0500	0.0000	0.0000	0.0000
SP_Bonds	LogisticRegression	0.9880	0.0500	0.0118	0.2000	0.0222
SP_Bonds	XGBoost	0.9880	0.0500	0.0333	0.2000	0.0571
SP_Only	LogisticRegression	0.9759	0.0286	0.0100	0.2000	0.0190
SP_Only	XGBoost	0.9880	0.0500	0.0333	0.2000	0.0571

Received: August 03,2025

International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The second robustness experiment examined label sensitivity, that is, how changes in crisis definitions (thresholds t1 and t2) and forecast horizons (5, 20, 60 business days) affected performance. As expected, less stringent thresholds and longer forecast horizons increased the number of positive (crisis) instances, leading to improved recall but reduced precision. Conversely, stricter definitions made crises rarer and harder to predict but produced higher-confidence signals when they occurred. The models exhibited stable ROC-AUC values across all configurations, indicating consistent discrimination power regardless of label frequency. This robustness across multiple label schemes suggests that the core features, volatility, yield slopes, and correlations, carry universal predictive information about systemic risk, independent of how “crisis” is specifically defined. A third analysis compared rolling retraining against a fixed train-test split. In the fixed approach, models were trained on an early segment (e.g., 2010–2018) and evaluated on later data (2019–2025). This configuration resulted in noticeable performance decay, particularly for XGBoost, whose feature relationships drifted over time as market regimes evolved. In contrast, the rolling retrain method, where models were periodically updated with the latest available data, preserved accuracy and improved recall during the post-2020 regime shifts. This result highlights the necessity of adaptive learning frameworks in financial prediction, as market structure and investor behavior continuously evolve. Models that fail to retrain regularly risk becoming obsolete or biased toward past dynamics.

Table 2. Label sensitivity sweep results

Label / Model		Fold	ROC_AUC	PR_AUC	Precision	Recall	F1
composite_crisis_20d_t1	LogisticRegression	3.0000	0.9679	0.0222	0.0167	0.2000	0.0308
composite_crisis_20d_t1	XGBoost	3.0000	0.9880	0.0500	0.0000	0.0000	0.0000
composite_crisis_5d_t1	LogisticRegression	5.0000	NaN	0.0000	0.0000	0.0000	0.0000
composite_crisis_5d_t1	XGBoost	5.0000	NaN	0.0000	0.0000	0.0000	0.0000
composite_crisis_60d_t1	LogisticRegression	3.0000	0.8530	0.3079	0.0698	0.2200	0.1052
composite_crisis_60d_t1	XGBoost	3.0000	0.8533	0.4695	0.2000	0.2000	0.2000
composite_crisis_60d_t2	LogisticRegression	3.0000	0.9933	0.1111	0.0857	0.2000	0.1200

Received: August 03,2025

composite_crisis_60d_t2	XGBoost	3.0000	0.9960	0.1400	0.0857	0.2000	0.1200

Together, these robustness and ablation experiments confirm the generalizability and economic validity of the modeling framework. The improved performance from cross-market features validates the systemic nature of financial crises, while the label and retraining sensitivity tests ensure that results are not overfit to arbitrary definitions or static data regimes. By combining these checks with the explainability insights derived from SHAP analysis, the research establishes a methodologically sound and economically interpretable foundation for machine learning–based systemic risk forecasting. The models demonstrate both conceptual and practical robustness, making them credible candidates for deployment in real-time financial stability monitoring and decision-support systems.

5. Discussion and Insights

The findings from this study reveal that multi-market modeling substantially enhances the consistency and robustness of financial crisis detection within the U.S. context. By integrating information from the S&P 500, U.S. Treasury bond yields, and the EUR/USD exchange rate, the models captured systemic linkages that are often overlooked in traditional, single-market analyses. This integration allowed the framework to detect early signals of financial distress that manifested across asset classes rather than within one isolated market. The results reinforce the idea that financial crises are not confined to individual segments of the economy but rather arise from cross-market contagion mechanisms that transmit shocks across equities, bonds, and foreign exchange simultaneously. Zhang et al. (2023) demonstrated that global systemic risk is driven by dynamic interdependencies among these markets, with correlation networks tightening during stress periods such as the COVID-19 pandemic [26]. The present findings align with this evidence, showing that similar contagion patterns are observable within the U.S., especially during episodes like the 2020 pandemic and the 2022 monetary tightening cycle, where volatility in equities coincided with yield curve inversions and capital flows reflected in the EUR/USD exchange rate.

The results highlight that multi-market modeling not only improves predictive consistency but also enhances the interpretability of systemic financial stress. When analyzed independently, market indicators such as the S&P 500 volatility or Treasury yield slopes provide valuable but incomplete information. However, when these indicators are jointly modeled, the system can capture the transmission of risk between markets, how a surge in equity volatility might be associated with widening credit spreads or sharp movements in bond yields. This interdependence represents the core of modern financial instability. Chen et al. (2025) showed through risk contagion network modeling that understanding the directionality and magnitude of these linkages is crucial for identifying early signals of systemic crises [6]. The present analysis supports this conclusion in the U.S. setting, demonstrating that the inclusion of cross-market correlations and contagion features led to more timely and stable crisis signals. The

Received: August 03,2025

International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

rolling correlation features and PCA-derived systemic components consistently ranked among the most influential predictors in SHAP analyses, further emphasizing that crisis prediction must account for dynamic inter-market relationships rather than static, univariate patterns.

A major insight from this research is the continued dominance of volatility and yield curve measures as early-warning indicators of financial distress. Volatility, especially rolling 20-day and 60-day standard deviations of returns, emerged as one of the most powerful signals preceding major market downturns. This is consistent with financial theory, which associates heightened volatility with rising uncertainty, liquidity tightening, and investor risk aversion. The U.S. experience during periods such as 2011, 2018, and 2022 illustrates how volatility spikes often predate broader market corrections. Similarly, the yield curve slope, the difference between long-term and short-term Treasury yields, remains an essential component of systemic risk assessment. Inversions in the U.S. yield curve have historically preceded recessions and banking stress, signaling investor expectations of economic slowdown and tightening financial conditions. By quantifying yield curve steepness and changes in its slope, the model effectively captured macro-financial dynamics associated with impending crises. The SHAP analysis reinforced this point by assigning consistently high importance to yield curve–related features, confirming that traditional macro indicators retain predictive validity even in machine learning–driven frameworks.

Importantly, the odels developed in this study were not designed as black-box predictors but as data-driven representations of economic intuition. Logistic Regression, with its interpretable coefficients, provided clarity on how variables such as rolling volatility and yield slope directly influence crisis probabilities. Even for more complex models like XGBoost, explainability tools such as SHAP values allowed for the decomposition of predictions into interpretable contributions. This interpretability bridges the gap between quantitative modeling and financial theory, making the results actionable for regulators and institutional investors. As Jabeur et al. (2021) emphasized, interpretable AI techniques like SHAP and feature attribution frameworks are vital for ensuring that predictive models in finance can be trusted, audited, and aligned with human decision-making [15]. In this study, SHAP-based analysis not only identified dominant features but also revealed how their influence evolved, reflecting shifts in market regimes and investor sentiment. For instance, during 2020–2021, bond-related features carried stronger predictive weight, whereas in 2022–2023, equity volatility and cross-market contagion indicators became more prominent, mirroring real-world market dynamics.

Beyond interpretability, this study underscores a deeper conceptual insight: machine learning models in finance do not replace economic reasoning, they extend it. The predictive framework successfully encoded long-observed relationships such as the inverse link between bond yields and equity returns, the signaling power of volatility, and the contagion of risk across asset classes. Rather than functioning as purely statistical instruments, these models served as computational amplifiers of economic knowledge, capable of quantifying complex nonlinearities and temporal dependencies that traditional econometric tools struggle to capture. This fusion of empirical data and financial logic positions multi-market machine learning systems as valuable tools for early-warning systems, portfolio risk management, and

Received: August 03,2025

macroprudential oversight within the U.S. financial system. In essence, the discussion reveals that the fusion of market domains enhances crisis detection consistency, volatility, and yield dynamics remain foundational indicators, and interpretability ensures alignment between model predictions and financial intuition. As U.S. financial markets continue to evolve under growing complexity and interconnectedness, the integration of explainable machine learning with multi-market data provides a scalable, transparent, and economically grounded approach to systemic risk forecasting and early crisis identification.

6. Limitations and Future Work

While this study provides strong evidence supporting the feasibility and interpretability of a multi-market machine learning framework for U.S. financial crisis prediction, several limitations should be acknowledged to contextualize the findings and guide future research. A primary limitation is the restricted temporal coverage of the dataset, which spans from 2010 to 2025. This period, though rich in events such as the European sovereign debt crisis, the 2020 COVID-19 pandemic, and the 2022 monetary tightening cycle, does not encompass earlier systemic events like the 2008 Global Financial Crisis (GFC), the 2000 dot-com bubble, or the 1998 Russian default. As a result, the models were trained and evaluated within a regime of post-GFC market dynamics characterized by prolonged low interest rates, quantitative easing, and structural shifts in capital flow behavior. Excluding older crises limits the generalizability of the model to historical crisis typologies that emerged under different monetary and financial regimes. Future extensions should therefore include longer historical windows, integrating data from the 1980s and 1990s to test the model's robustness across multiple macro-financial cycles. This broader temporal scope would also allow for studying how structural transformations in the U.S. economy, such as deregulation, digitalization, and monetary policy evolution, affect the predictability and propagation of systemic risk.

Another limitation is the exclusion of macroeconomic, sentiment, and policy variables. While the study focused exclusively on market-based indicators (equities, bonds, and foreign exchange), financial crises are often preceded or amplified by shifts in economic fundamentals, such as credit growth, unemployment, and inflation expectations, as well as by shifts in investor sentiment captured through measures like news tone or social media sentiment indices. Integrating these dimensions could enrich the explanatory power of the models and align them more closely with the complex, multi-causal nature of financial crises. For instance, incorporating macroeconomic signals like credit-to-GDP gaps, monetary policy stance, or consumer confidence indices could help differentiate between liquidity-driven volatility and systemic distress. Similarly, sentiment analysis using text-based data could provide an early-warning component rooted in behavioral finance, capturing changes in market psychology that precede asset repricing events.

From a methodological standpoint, the study relied primarily on logistic regression and XGBoost as the core predictive models. While these models offer interpretability and robustness, they are inherently limited in capturing long-term temporal dependencies and complex nonlinear feedback loops that characterize financial contagion. Future research should explore advanced deep learning architectures, particularly Long Short-Term Memory (LSTM)

Received: August 03,2025

International Journal of Applied Mathematics
Volume 38 No. 8s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

and Transformer-based models, which have demonstrated strong capabilities in sequence modeling and multivariate time-series forecasting. Tölö (2020) showed that recurrent neural networks can effectively capture temporal dependencies in systemic financial risk, outperforming traditional econometric methods when dealing with high-dimensional, sequential data [24]. Similarly, Kohan Sefidi (2025) employed LSTM-based multilayer network analysis to model spillover effects between financial markets, demonstrating improved performance in identifying crisis propagation patterns and inter-market linkages [17]. Adapting these architectures to a multi-market setting could enable the detection of lagged contagion effects, how shocks in one asset class gradually permeate others, thus refining the model's temporal precision and predictive stability.

A further avenue for development lies in the incorporation of causal inference and explainable policy modeling. While this study's use of SHAP values provided transparency regarding feature importance, it did not establish causal relationships between indicators and crisis outcomes. Causal machine learning techniques, such as causal forests, instrumental variable regression with high-dimensional controls, or structural equation modeling, could bridge this gap by distinguishing between predictive associations and true causal drivers. Such an approach would be particularly valuable for policy applications, where regulators require interpretable, causally grounded insights to guide interventions such as capital buffer adjustments, liquidity provisions, or macroprudential policy responses. Finally, expanding the research beyond the U.S. to a global multi-market framework could substantially enhance understanding of international spillovers and contagion dynamics. Integrating datasets from major economies such as the Eurozone, Japan, and emerging markets would allow for comparative studies of crisis propagation, providing insights into how global shocks, such as energy price surges or geopolitical disruptions, affect interconnected financial systems. A cross-country model could also explore how exchange rate regimes, capital flow controls, and fiscal policies mediate the transmission of financial stress across borders. In conclusion, while this research advances the state of machine learning–based financial crisis prediction by integrating U.S. equity, bond, and forex markets into a unified framework, its limitations highlight opportunities for richer, more holistic modeling. Future work should combine deep neural architectures, macroeconomic and sentiment indicators, causal inference, and global market extensions to build a comprehensive early-warning system capable of supporting both investors and policymakers in managing systemic financial risk.

7. Conclusion

Summarize that integrating multi-market data improves systemic risk prediction and interpretability.

Highlight model transparency, robust validation, and economic interpretability as key advances toward data-driven financial crisis monitoring frameworks in the U.S.

Received: August 03,2025

References

[1] Ahad, M. A., et al. (2025). AI-Based Product Clustering for E-Commerce Platforms: Enhancing Navigation and User Personalization. International Journal of Environmental Sciences, 156–171.

[2] Benzoni, L., & Chyruk, O. (2018). Why does the yield-curve slope predict recessions? Chicago Fed Letter, 404, 1–8.

[3] Bluwstein, K., Buckmann, M., Joseph, A., Kang, M., Kapadia, S., & Şimşek, Ö. (2020). Credit growth, the yield curve and financial crisis prediction: evidence from a machine learning approach. Bank of England Staff Working Paper No. 848.

[4] Campisi, G., Muzzioli, S., & De Baets, B. (2024). A comparison of machine learning methods for predicting the direction of the U.S. stock market. Journal of Forecasting, 43(1), 89–

[5] Casabianca, E. J., Catalano, M., Forni, L., Giarda, E., & Passeri, S. (2022). A machine learning approach to rank the determinants of banking crises over time and across countries. Journal of International Money and Finance, 129, 102739.

[6] Chen, B., Li, S., & Wang, H. (2025). Risk contagion network and characteristic measurement across markets. Journal of Financial Markets, 58(4), 45–67.

[7] Chen, Y., Zhang, L., & Wang, H. (2024). Forecasting stock market crisis events using machine learning. Expert Systems with Applications, 198, 116847.

[8] Das, B. C., Sartaz, M. S., Reza, S. A., Hossain, A., Nasiruddin, M. D., Bishnu, K. K., ... & Abed, J. (2025). AI-Driven Cybersecurity Threat Detection: Building Resilient Defense Systems Using Predictive Analytics. arXiv preprint arXiv:2508.01422.

[9] Ding, H., Zhao, X., Jiang, Z., Abdullah, S. N., & Dewi, D. A. (2024). EUR/USD exchange rate forecasting based on information fusion with large language models and deep learning methods. Expert Systems with Applications, 224, 120019.

[10] Fariha, N., et al. (2025). Advanced fraud detection using machine learning models: Enhancing financial transaction security. arXiv preprint arXiv:2506.10842.

[11] Fouliard, J., Howell, M., Rey, H., & Stavrakeva, V. (2021). Answering the Queen: Machine learning and financial crises. NBER Working Paper No. 28302. National Bureau of Economic Research.

[12] Gholkar, R., & Sueppel, R. (2024). Using principal components to construct macro trading signals. Quantitative Finance, 24(8), 1123–1142.

[13] Gieck, J. (2012). Interdependence and contagion in global asset markets. ECB Working Paper Series No. 1480. European Central Bank.

[14] Islam, M. Z., Rahman, M. S., Sumsuzoha, M., Sarker, B., Islam, M. R., Alam, M., & Shil, S. K. (2025). Cryptocurrency Price Forecasting Using Machine Learning: Building Intelligent Financial Prediction Models. arXiv preprint arXiv:2508.01419.

[15] Jabeur, S. B., Gharib, C., Mefteh-Wali, S., & Arfi, W. B. (2021). CatBoost model and artificial intelligence techniques for corporate failure prediction. Technological Forecasting and Social Change, 166, 120658.

Received: August 03,2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[16] Khan, M. N. M., et al. (2025). Assessing the Impact of ESG Factors on Financial Performance Using an AI-Enabled Predictive Model. International Journal of Environmental Sciences, 1792–1811.

[17] Kohan Sefidi, M. (2025). Predicting financial market crises using multilayer network analysis and LSTM-based forecasting of spillover effects. Journal of Financial Stability, 62, 101245.

[18] Marathe, S. D. (2024). Demystifying interpretable AI in finance: A review of SHAP and LIME. International Journal of Innovative Research in Technology, 11(5), 1633–1641.

[19] Reza, S. A., Rahman, M. K., Hossain, M. S., Rabbi, M. N. S., Quddus, A., Mozumder, S. S., ... & Ahad, M. A. (2025). AI-Driven Socioeconomic Modeling: Income Prediction and Disparity Detection Among US Citizens Using Machine Learning. Advances in Consumer Research, 2(4).

[20] Roszyk, N., & Ślepaczuk, R. (2024). The hybrid forecast of S&P 500 volatility ensembled from VIX, GARCH and LSTM models. Journal of Forecasting, 43(2), 341–365.

[21] Schnaubelt, M. (2019). A comparison of machine learning model validation schemes for non-stationary time series data. Journal of Financial Data Science, 1(2), 107–130.

[22] Shawon, R. E. R., Supti, A. M., Kawsar, M., Al Jobaer, M. A., Islam, M. F., Sumon, P. D., ... & Malla, S. (2025). Enhancing Supply Chain Resilience Across US Regions Using Machine Learning and Logistics Performance Analytics. International Journal of Applied Mathematics, 38(4s).

[23] Sizan, M. M. H., Chouksey, A., Dola, A., Begum, S., Antara, U. K., Sazid, F., & Oishi, R. A. (2025). Machine Learning-Based Unsupervised Ensemble Approach for Detecting New Money Laundering Typologies in Transaction Graphs. International Journal of Applied Mathematics, 38(2s).

[24] Tölö, E. (2020). Predicting systemic financial crises with recurrent neural networks. Journal of Financial Stability, 49, 100756.

[25] Wang, C. S. H., Liu, M., & Zhang, H. (2023). Market systemic risk, predictability and macroeconomics. Journal of Financial Economics, 148(2), 125–147.

[26] Zhang, P., Wang, L., & Chen, Y. (2023). Global systemic risk dynamic network connectedness during COVID-19: Evidence from stock, bond, and foreign exchange markets. International Review of Economics & Finance, 87, 240–258.

[27] Zhao, X., Huang, C., Yang, X., Cao, J., & Yang, X. (2024). Can we better predict financial crisis? The role of Laplacian-energy-like measure. International Review of Financial Analysis, 91, 103045.

Received: August 03,2025