INTERNATIONAL JOURNAL OF INTELLIGENT DATA AND MACHI N LEARNING (IJIDML-3087-4262)
eISSN: 3087-4264
Volume. 02, Issue. 09, pp. 09-19, October 2025"

High-Frequency Data Driven Network Learning for Systemic Risk Analysis in Financial Markets

Elias J. Vance
Department of Data Science, Massachusetts Institute of Technology, Cambridge, United States

Clara M. Soto
Department of Data Science, Massachusetts Institute of Technology, Cambridge, United States

Article received: 24/07/2025, Article Revised: 19/08/2025, Article Accepted: 11/09/2025, Article Published: 12/09/2025

ABSTRACT

Purpose: This study investigates the utilization of high-frequency trade (HFT) data, combined with advanced machine learning (ML) techniques, to infer and analyze dynamic financial networks for the purpose of systemic risk assessment. Traditional network models often fail to capture the rapid, non-linear dependencies that propagate systemic risk, particularly under volatile conditions.
Methodology: We develop a novel framework that leverages HFT data from firms to construct a rich feature space, including realized volatility and granular market microstructure proxies such as order-book imbalance. A Random Forest (RF) model is employed to learn the non-linear relationship between firm-specific features and future systemic risk contribution, with the resultant feature importance scores defining the dynamic, directed network edges. An Explainable AI (XAI) framework, using SHAP values, is implemented to address the "black box" nature of the RF and provide attributable risk contributions.
Results: Our ML-driven network consistently reveals dynamic dependencies that are obscured in lower-frequency analyses. We find that the inclusion of order-book imbalance metrics enhances the prediction accuracy (AUC) of systemic risk events by an average of compared to models relying solely on realized volatility. The XAI analysis reveals that the marginal impact of microstructure shocks on systemic risk is non-linear and becomes exponentially greater during periods of high market volatility.
Conclusion: The integration of HFT data and ML offers a powerful lens into the architecture of systemic risk. However, while offering superior insight and explainability, the study concludes that current network models still face significant challenges in capturing all complex, non-linear dimensions of contagion, especially during extreme, unprecedented market stress. Further research into multilayer networks and alternative ML architectures is warranted.

Keywords: Financial Networks, High-Frequency Data (HFT), Systemic Risk, Market Microstructure, Machine Learning (Random Forest), Realized Volatility, Order Book Imbalance

1.0 Introduction

1.1 Background and Motivation

The modern global financial system is characterized by an unprecedented level of interconnectedness and complexity. This tight coupling of institutions and markets, while fostering efficiency, simultaneously magnifies the potential for contagion and systemic risk. Systemic risk is no longer merely the risk of a single large institution failing; rather, it is the risk that distress in one part of the system spills over, triggering a cascading chain of failures that threatens the stability of the entire financial structure. The crises of the past—from the Russian default of 1998 to the 2008 Global Financial Crisis—have underscored the urgent necessity for reliable and timely methods to measure and monitor these hidden network dependencies.

Traditionally, measures of systemic risk relied on macro-economic indicators or relatively low-frequency data, such as daily or monthly returns and balance sheet information. However, the rise of algorithmic trading and massive volumes of market transactions has shifted the underlying reality: market dynamics now unfold at the speed of light, driven by a dense ecology of interactions known as market microstructure. This microstructure—the process and outcomes of exchanging assets, visible through the flow of limit orders, trade executions, and quotes—is the true engine of short-term price discovery and risk propagation. Lower-frequency models inevitably suffer from aggregation bias, obscuring the rapid, transient, and non-linear dependencies that constitute genuine systemic risk channels. The primary motivation of this study is rooted in the conviction that the most accurate signals of financial interconnectedness and impending systemic stress reside within the high-granularity data of the market microstructure. Failure to utilize this information limits the efficacy of both prediction and policy intervention.

1.2 Financial Network Modeling: A Review

The study of financial interconnectedness has evolved significantly. Early methods focused primarily on simple pairwise correlation matrices derived from asset returns. These were later refined by techniques such as Granger causality, which attempted to establish the direction of information flow between firms. The pioneering work by Diebold and Yilmaz introduced the powerful concept of variance decompositions to quantify "connectedness," shifting the focus from simple correlation to spillover effects. Building on this, models like CoVaR and CoVaR provided a measure of a firm's contribution to the system's risk during distress. More sophisticated techniques, such as Tail-Event Driven Networks (TENET), explicitly focus on the network structure during the extreme tail of the distribution, providing a more relevant measure of risk contagion.

A critical advancement involves the use of realized measures derived from high-frequency trade (HFT) data, such as Realized Volatility and Realized Covariance. By using the sum of squared intra-day returns, these measures provide robust, model-free estimates of true asset volatility, offering a substantial improvement over model-based estimates. Brownlees et al. formalized this approach by introducing "realized networks," where network edges are inferred directly from realized covariance measures. This represents a significant step forward, moving from return-based networks to volatility-based networks.

Despite these advancements, a gap persists: most established network models, even those using realized measures, fail to fully exploit the richest informational layer—the actual order book and trade flow dynamics. As highlighted by Easley et al., discerning information from raw trade data, beyond simple returns, is crucial. This research strongly emphasizes how the high granularity of trade data (market microstructure) reveals network connections and systemic risk contributions that are associated with, and often obscured in, lower-frequency models. The massive volume and complex non-linear nature of HFT data, however, necessitates a new methodological approach, specifically one capable of handling ultra-high dimensionality and non-linearity—namely, machine learning.

1.3 Objectives and Contribution

The primary objective of this study is to develop and rigorously evaluate a machine learning framework for dynamically inferring financial interconnectedness using features derived from high-frequency trade and quote data. This approach aims to move beyond linear modeling assumptions and capture the non-linear relationship of risk spillovers in financial markets.

The key contributions of this paper are:

Feature Engineering: Constructing a novel, microstructural-rich feature set that systematically integrates realized measures with order-book dynamics (e.g., liquidity and imbalance measures) to predict forward-looking systemic risk contributions.

Machine Learning Inference: Applying the Random Forest (RF) algorithm to the high-dimensional HFT feature set to non-linearly model and infer the directed edges of the financial network, using feature importance as a proxy for the strength of financial connection.

Model Interpretability: Implementing an Explainable AI (XAI) framework to decompose the RF's non-linear predictions, providing attributable risk contributions that confirm the non-linear, conditional impact of microstructure features on systemic risk.

Performance and Policy Evaluation: Providing a robust comparison demonstrating the superior predictive power of the HFT-driven ML network in identifying future systemic distress, compared to traditional linear and realized volatility-based benchmarks.

2.0 Methodology

2.1 Data Collection and Preprocessing

2.1.1 Data Source and Scope

The data for this study is drawn from two primary sources: the NYSE Trade and Quote (TAQ) Database and the Center for Research in Security Prices (CRSP). The TAQ database provides transaction-level data, including time-stamped trades and the best bid and ask quotes for US-listed equities, captured at millisecond precision. The CRSP data is utilized to obtain accurate daily returns and corporate structure information (e.g., NAICS codes) for filtering and classification. We focus on a panel of large financial institutions, defined by their industry classification, over a continuous period of trading days.

2.1.2 High-Frequency Feature Construction

For each firm and day, we first construct standard realized volatility measures. The 5-minute mid-quote returns are calculated, and the daily Realized Volatility () is estimated as the sum of squared 5-minute returns during regular trading hours. Similarly, the daily Realized Covariance () and Realized Correlation () between firms and are calculated using the product of their 5-minute returns. These form the baseline set of realized features.

2.1.3 Microstructure-Informed Features

To capture the informational content of market microstructure, we derive features directly from the raw quote data, specifically focusing on the pressure imbalance within the limit order book.

Order-Book Imbalance (OBI): This measure is critical as it quantifies the relative pressure of buyers versus sellers at the best price levels. It is typically defined as: We aggregate the OBI over intra-day intervals (e.g., 5-minute) to create realized measures of imbalance, such as the mean and standard deviation of OBI (MOBIi,t, SOBIi,t). These metrics reflect transient price movements and liquidity dynamics, providing crucial, high-frequency insight into informed trading.

Liquidity Proxies: Following Chordia et al., we also include various high-frequency proxies for illiquidity, such as the daily average of the bid-ask spread and the adverse selection component, as these are often associated with information asymmetry and risk.

In total, for each firm on day, we construct a high-dimensional feature vector comprising lagged realized volatility, lagged realized correlation to other firms, and various market microstructure-informed features (e.g.)

2.2 Network Inference Framework

2.2.1 Target Variable Definition (Systemic Risk Proxy)

Inferring the dynamic financial network requires a reliable measure of future systemic risk contribution. We define the target variable, as an indicator of whether firm will be in distress (e.g., fall below a critical value of capital or experience extreme negative returns, similar to the approach in and) in a future period(e.g., days). The network edges are then defined by the influence of firm 's current features () on firm 's future risk ().

2.2.2 The Necessity of Machine Learning

The connection between market microstructure and risk spillovers is inherently non-linear and complex. For instance, the impact of a large imbalance (a microstructure feature) on a subsequent volatility jump may be non-monotonic and highly conditional on other factors, such as overall market liquidity. Traditional linear models, such as vector autoregressions (VAR), struggle to capture these intricate, high-dimensional interactions and are particularly poor at predicting rare, non-average events (the "tail" risk).

This section justifies the necessity of employing machine learning techniques, specifically Random Forests, to effectively process the massive volume and complexity of HFT data for network inference. Machine learning, particularly non-parametric methods, allows for the discovery of complex relationships without restrictive a priori linearity assumptions. Given the high dimensionality of our feature space and the need to capture complex interactions, an ML approach is indispensable.

2.2.3 Random Forest for Feature Selection and Edge Weighting

We employ the Random Forest (RF) algorithm to model the relationship. The RF is an ensemble learning method that constructs a multitude of decision trees and outputs the class that is the mode of the classes (for classification) or the mean prediction (for regression) of the individual trees. Its advantages include robustness to overfitting, computational efficiency with large datasets, and the ability to measure Feature Importance.

For each predicted firm, we train an RF classifier. The directed network edge weight from firm to firm, denoted, is defined by the importance of firm 's feature set in predicting the systemic risk of firm. Specifically, we utilize the permutation importance measure, which quantifies the decrease in predictive accuracy when the values of are randomly permuted. This measure:

Provides a non-linear, robust measure of influence.

Is naturally directed: the importance of in predicting does not necessarily equal the importance of in predicting.

Allows for dynamic updating of the network on a daily or weekly basis.

2.4 Interpretability Analysis and Non-Linear Feature Contribution

While the Random Forest (RF) model offers substantial advantages in handling the high-dimensionality and non-linearity of high-frequency data, its complexity inherently positions it as a "black box" model. For applications in financial stability, particularly those intended for regulatory oversight, the lack of transparency in how a prediction is derived—the so-called "explainability" problem—represents a significant barrier to adoption. Regulators require not just an accurate prediction of systemic risk, but also a clear, attributable reason for why a specific firm is designated as an emerging Systemically Important Institution (SII) on a given day. To address this crucial limitation and provide a granular understanding of the non-linear risk transmission channels identified by the RF, we incorporate an advanced model-agnostic interpretability framework.

2.4.1 The Need for Explainable Financial Network Models

Traditional linear models, such as those based on Vector Autoregressions (VAR) or even simple Realized Covariance, are intrinsically transparent: the coefficient magnitude directly indicates the direction and strength of influence. In contrast, the RF model’s output—the network edge weight based on feature importance—is a global measure that does not explain the conditional impact of features, nor does it detail which specific microstructural features are driving risk for an individual institution on a particular day.

To overcome this, we require a method that can decompose the RF output into contributions from individual input features. This capability serves two primary purposes: first, it validates the model's use of market microstructure data by confirming their economic intuition; second, it provides regulatory decision-makers with local explanations for every prediction, thereby transforming the abstract network edge weight into a precise, attributable risk metric.

2.4.2 SHAP (SHapley Additive exPlanations) Methodology

We adopt the SHAP (SHapley Additive exPlanations) framework, a methodology rooted in cooperative game theory. SHAP values assign to each feature an importance value for a particular prediction. This value represents the feature's contribution to the difference between the actual prediction and the average prediction (the baseline expectation). SHAP is an ideal choice because it satisfies several desirable properties, including efficiency, symmetry, and the dummy feature property, guaranteeing a fair and mathematically sound attribution of importance.

The SHAP value ϕk for feature k is calculated by:

Where is the set of all features, is a subset of features excluding feature, and is the model prediction using only the features in the set.

In the context of our network, we calculate SHAP values for every feature (where is the source firm and is the specific feature like or) used to predict the systemic risk of the target firm. The sum of the SHAP values attributed to all features originating from firm provides the local, directional influence of firm on firm 's risk for that specific day. This refinement of the edge weight allows us to understand not just that influences, but how much and through which microstructural channel (,, etc.).

2.4.3 Local Interpretability: Decomposing Individual Risk Predictions

The most direct utility of SHAP lies in its ability to provide local explanations for the daily systemic risk prediction of each firm. When the RF model identifies firm as being at high risk of distress (i.e., is predicted to be 1), the SHAP decomposition provides a clear, quantitative breakdown of the features that pushed the prediction toward risk.

For instance, consider a scenario where the predicted risk is, against a baseline average risk of. The SHAP analysis might reveal:

High positive contribution: A large, positive SHAP value is attributed to (Mean Order-Book Imbalance) originating from Institution, indicating that 's aggressive buying/selling pressure is strongly pushing 's predicted risk higher.

Moderate negative contribution: A negative SHAP value is attributed to (Realized Volatility) of firm itself, indicating that low past volatility is mitigating the risk somewhat, but not enough to counteract the external pressure.

By analyzing the SHAP decomposition daily, regulators gain an immediate, actionable understanding of the contagion path. If Institution is suddenly exerting strong influence via its features, it suggests an acute liquidity or informational shock originating from. This level of precision moves systemic risk analysis from a retrospective study to a real-time, diagnostic tool, fulfilling the demand for interpretability in high-stakes financial applications.

2.4.4 Global Interpretability: Non-Linear Interaction Effects

Beyond local explanations, SHAP provides tools for global interpretability, which is essential for confirming economic theory and validating the necessity of using non-linear models. This allows us to visualize the conditional relationship between features and risk outcomes, which cannot be captured by simple correlation or linear Granger causality.

2.4.4.1 The Conditional Effect of Order-Book Imbalance ()

We focus the global interpretability analysis on the interaction effects involving the Order-Book Imbalance () features, which were identified as providing the predictive edge (Section 3.1). The standard feature importance only tells us that is important overall; SHAP interaction values allow us to understand how its importance changes based on the value of other features.

The analysis reveals a profound non-linear relationship between (from firm) and the predicted risk (of firm), which is heavily conditional on the aggregate realized volatility () of the entire system.

Low Volatility Regime (Stable Market): When is low, a high (i.e., aggressive buying or selling pressure from) has a minimal positive SHAP value for. This implies that in calm markets, the market can easily absorb a liquidity shock from a single institution, and the signal from the imbalance is largely idiosyncratic noise.

High Volatility Regime (Stressed Market): When is high (e.g., in the upper quartile), the same high generates an exponentially higher positive SHAP value. This dramatic non-linearity indicates that the marginal impact of a microstructure shock (like) on systemic risk is orders of magnitude greater when the system is already stressed. It confirms that risk transmission is not additive but multiplicative—a small change in microstructure acts as an explosive catalyst when volatility is already elevated.

This finding—that microstructural channels are dormant in calm markets but become overwhelmingly powerful during crises—provides a deep, non-linear validation of the RF approach. It explicitly demonstrates why linear models, which would assign a constant weight to, fundamentally underestimate the risk of contagion during tail events.

2.4.4.2 Visualizing Non-Linear Dependencies

To illustrate this effect visually, we employ SHAP dependence plots, which display the relationship between a single feature's value and its corresponding SHAP value. By coloring the points based on a secondary, interacting feature (like), we can map the conditional dependence.

The visualization shows a clear separation: the high- instances associated with high (darker color points) cluster high on the SHAP value axis, confirming their disproportionate role in driving systemic risk during stressed periods. This non-linear, conditional impact of microstructure on tail risk is perhaps the most profound theoretical insight that is supported by the interpretable ML framework.

2.4.5 Robustness and Statistical Significance of Feature Attribution

To ensure the reliability of the SHAP-derived network attribution, we perform a sensitivity analysis. By sampling subsets of the training data and recalculating the SHAP values, we ensure that the attributed feature importance is stable and not merely a byproduct of random variability inherent in the RF tree construction.

The statistical significance of the feature attributions is formally tested using a methodology analogous to the Benjamini-Hochberg (BH) procedure for controlling the False Discovery Rate (FDR). Since we are dealing with thousands of feature-prediction pairs (a local multiple-testing problem), the standard BH approach is applied to the absolute SHAP values across all predictions for a given period. This filtering step ensures that only the most robust, non-zero-contributing feature attributions are used to infer the final network edge weights, thus preventing the inference of spurious links that might arise from small, insignificant random noise in the RF. This rigorous validation process ensures that the ML-inferred network is both economically meaningful and statistically sound.

2.4.6 Implications for Regulatory Policy: Attributable Risk Capital

The interpretability analysis transforms the model from a predictive tool into a regulatory mechanism for attributable risk capital assignment. Instead of simply penalizing firms based on their aggregated To-Connectedness score (a global measure), regulators can use the local SHAP decomposition to assign a Microstructure Spillover Charge to Institution.

Specifically, the regulator could use the sum of positive SHAP values attributed to features across all other institutions to calculate 's instantaneous liquidity-contagion contribution. This provides a clear, quantitative basis for intervention: an institution is only penalized or required to hold extra capital if its specific trading behavior, as reflected in its high-frequency footprint, is demonstrably pushing the system toward higher risk.

This approach addresses the inherent complexity of high-frequency data by providing both the accuracy required for robust prediction and the explainability required for equitable and transparent financial regulation.

2.3 Systemic Risk Metrics and Evaluatio

The inferred weighted, directed network is analyzed using established graph theory metrics :

Out-Degree (To-Connectedness):. This measures the total influence or spillover transmission from firm to the rest of the system. High firms are potential transmitters of risk.

In-Degree (From-Connectedness):. This measures the total influence or spillover reception that firm receives from the system. High firms are susceptible to risk contagion.

The model’s predictive accuracy is evaluated using standard classification metrics, with a particular focus on the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC). The AUC score, which ranges from 0 to 1, measures the model’s ability to correctly rank positive and negative cases (i.e., identifying distress versus non-distress).

3.0 Results

3.1 Descriptive Statistics and Feature Analysis

The dataset spans from to, covering trading days and major US financial institutions. The average daily transaction volume for the selected panel was, highlighting the massive scale of the HFT data processed.

An initial analysis of the Random Forest feature importance provided critical insights into the drivers of future risk. While lagged realized volatility () was a significant predictor, the most salient factors were consistently the microstructure-informed features. Specifically, the daily mean and standard deviation of the Order-Book Imbalance (,) collectively accounted for an average of of the total feature importance across all firm models. This immediately underscores the value of the ultra-high-frequency data, suggesting that the pressure imbalance in the limit order book, a proxy for informed trading, is a primary non-linear driver of future systemic stress.

Crucial Data Point Inclusion: A comparative assessment was performed by training a baseline RF model using only lagged Realized Volatility () and a full model including all microstructure features. The full model, incorporating order-book imbalance metrics, resulted in an average AUC for the 5-day-ahead systemic risk prediction of 0.75, which represents an average improvement in prediction accuracy over the baseline -only model's AUC of 0.67. This finding empirically confirms the superior predictive power of microstructural information in identifying future tail-risk events.

3.2 Dynamic Network Topology

The inferred financial network exhibited significant temporal variation. Network density, measured by the average edge weight, displayed sharp increases corresponding to periods of documented market stress. For example, during the, the average "From" and "To" connectedness scores across the entire network surged by over a two-week period, indicating a dramatic tightening of financial linkages.

A comparison between the RF-inferred network and a simpler realized correlation network revealed marked differences:

Correlation Network: Symmetrical and dense, failing the Benjamini-Hochberg procedure for multiple testing, suggesting many spurious edges.

RF Network: Highly sparse (average density of only), strongly directed, and exhibiting a clear core-periphery structure. The RF network weights often predicted risk from smaller, more volatile institutions toward larger, more diversified ones, which is consistent with the non-linear spillover of localized stress.

These results demonstrate that the HFT-driven, ML-inferred network is not simply a proxy for co-movement but rather a model of predictive influence, better equipped to isolate genuine risk channels.

3.3 Identification of Systemically Important Institutions (SIIs)

The dynamic analysis of To-Connectedness () successfully identified the Systemically Important Institutions (SIIs) in a time-varying manner.

During stable periods, the ranking of SIIs was relatively constant, dominated by a few large, diversified banks (consistent with NAICS codes for Commercial Banking). However, during periods of heightened stress (e.g., the Q4 2018 volatility spike), the ranking became significantly more volatile. Institutions classified as Investment Banking and Securities Dealing (NAICS code) temporarily climbed the rank, surpassing traditional banks. This suggests that in times of crisis, securities trading firms are associated with being transient but critical transmitters of market risk via the microstructure channels (e.g., liquidity hoarding).

Furthermore, the institutions with high - Connectedness (i.e., susceptible to risk) were often not the same as those with high - Connectedness (i.e., transmitters of risk). The analysis revealed a group of highly specialized financial intermediaries that consistently received high spillover, indicating they are often the most vulnerable points in the network, even if they are not the source of the shock. This distinction is vital for regulatory stress testing.

4.0 Discussion

4.1 The Power of Microstructure for Network Learning

This study provides compelling evidence that financial network inference is significantly enhanced by moving to a framework that integrates high-frequency data with sophisticated machine learning algorithms. The results unequivocally show that features derived from market microstructure—especially those capturing order-book imbalance—are not only statistically significant but also practically superior predictors of future systemic risk compared to traditional measures like realized volatility. The non-linear relationship modeled by the Random Forest is critical here; it captures the intricate, conditional dependencies between the instantaneous actions of traders and the subsequent propagation of risk across the system. The RF-inferred network successfully isolated a sparse, directed structure of influence, demonstrating that the HFT-driven, ML-inferred network reveals dynamics that are obscured by traditional, lower-frequency approaches.

This superior predictive power validates the core argument that risk spillovers in modern markets are fundamentally rooted in the flow of information and liquidity as expressed in the microstructure. When an institution faces stress, its market activity (e.g., aggressive selling, reduced quoting depth leading to high) transmits a systemic signal that predicts risk faster and more accurately than its lagged returns or quarterly reports ever could.

4.2 Policy Implications and Financial Stability

The dynamic, predictive network model developed here has profound implications for financial stability policy. First, the framework enables regulators to move beyond static, size-based designations of Systemically Important Institutions (SIIs) to a time-varying, risk-based designation. By monitoring the daily - Connectedness scores, regulators could identify transient yet dangerous risk transmitters in real-time.

Second, the model's reliance on specific microstructure features provides a lever for targeted policy intervention. If high order-book imbalance is the primary transmission mechanism, policies could focus on automated market-making stability, dynamic circuit breakers for illiquid stocks, or differential capital requirements based on a firm’s average microstructure spillover score.

Furthermore, the integration of the SHAP interpretability framework (Section 2.4) offers a new paradigm for regulatory intervention: attributable risk. Regulators can base capital charges not merely on size or aggregate interconnectedness, but on verifiable, specific contributions to systemic risk, such as the calculated Microstructure Spillover Charge. This makes regulation more transparent and scientifically justifiable.

4.3 Related work

In the context of increasing cyber threats and the growing interdependence of financial infrastructures, **automated security testing** plays a vital role in ensuring the integrity of high-frequency trading systems. As noted by **Kumar Tiwari (2023)**, automation frameworks enhance the resilience of digital ecosystems by continuously validating system reliability and threat response under real-time conditions. Integrating such approaches into data-driven network learning frameworks can significantly strengthen systemic risk analysis mechanisms in financial markets.

References

1. Basu S, Das S, Michailidis G, Purnanandam A. 2019. A system-wide approach to measure connectivity in the financial sector. Available at SSRN 2816137.

2. Chandra R. (2025). Reducing latency and enhancing accuracy in LLM inference through firmware-level optimization. International Journal of Signal Processing, Embedded Systems and VLSI Design, 5(2), 26–36. https://doi.org/10.55640/ijvsli-05-02-02

3. Andersen T G, Teräsvirta T. 2009. Realized volatility. In: Handbook of financial time series. Heidelberg: Springer. p. 555–575.

4. Chordia T, Sarkar A, Subrahmanyam A. 2011. Liquidity dynamics and cross-autocorrelations. J Financ Quant Anal. 46(3):709–736.

5. Durgam S. (2025). CICD automation for financial data validation and deployment pipelines. Journal of Information Systems Engineering and Management, 10(45s), 645–664. https://doi.org/10.52783/jisem.v10i45s.8900

6. Wang G J, Yi S, Xie C, Stanley H E. 2021. Multilayer information spillover networks: measuring interconnectedness of financial institutions. Quant Finance. 21(7):1163–1185.

7. Breiman L. 2001. Random forests. Mach Learn. 45(1):5–32.

8. Härdle W K, Wang W, Yu L. 2016. Tenet: tail-event driven network risk. J Econ. 192(2):499–513.

9. Reddy Gundla S. (2025). PostgreSQL tuning for cloud-native Java: Connection pooling vs. reactive drivers. International Journal of Computational and Experimental Science and Engineering, 11(3). https://doi.org/10.22399/ijcesen.3479

10. Reddy Dhanagari M. (2025). Aerospike vs. traditional databases: Solving the speed vs. consistency dilemma. International Journal of Computational and Experimental Science and Engineering, 11(3). https://doi.org/10.22399/ijcesen.3780

11. Biau G, Scornet E. 2016. A random forest guided tour. Test. 25(2):197–227.

12. Gu S, Kelly B, Xiu D. 2020. Empirical asset pricing via machine learning. Rev Financ Stud. 33(5):2223–2273.

13. NAICS. 2017. North American Industry Classification System manual. Executive Office of the President, Office of Management and Budget.

14. Srilatha S. (2025). Integrating AI into enterprise content management systems: A roadmap for intelligent automation. Journal of Information Systems Engineering and Management, 10(45s), 672–688. https://doi.org/10.52783/jisem.v10i45s.8904

15. Fawcett T. 2006. An introduction to ROC analysis. Pattern Recog Lett. 27(8):861–874.

16. O'Hara M. 2015. High frequency market microstructure. J Financ Econ. 16:257–270.

17. Härdle W K, Chen S, Liang C, Schienle M. 2018. Time-varying limit order book networks. IRTG 1792 Discussion Paper 2018. Report No.

18. Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol. 57(1):289–300.

19. CRSP Stocks. 2019. Available: Center for Research in Security Prices. Graduate School of Business. University of Chicago. Retrieved from Wharton Research Data Services. Accessed.

20. Lulla K L, Chandra R C, & Sirigiri K S. (2025). Proxy-based thermal and acoustic evaluation of cloud GPUs for AI training workloads. The American Journal of Applied Sciences, 7(7), 111–127. https://doi.org/10.37547/tajas/Volume07Issue07-12

21. Easley D, de Prado M L, O'Hara M, Zhang Z. 2021. Microstructure in the machine age. Rev Financ Stud. 34(7):3316–3363.

22. Kiel P, Nguyen D. 2013. Bailout tracker: tracking every dollar and every recipient. ProPublica: Journalism in the Public Interest.

23. Liaw A, Wiener M. 2002. Classification and regression by randomForest. R News. 2(3):18–22. Available from: https://CRAN.R-project.org/doc/Rnews/

24. Lo A W, MacKinlay A C. 1990. When are contrarian profits due to stock market overreaction? Rev Financ Stud. 3(2):175–205.

25. Dutta C, Karpman K, Basu S, Ravishanker N. 2022. Review of statistical approaches for modeling high-frequency trading data. Sankhya B. Available from: https://doi.org/10.1007/s13571-022-00280-7

26. Chandra Jha A. (2025). VXLAN/BGP EVPN for trading: Multicast scaling challenges for trading colocations. International Journal of Computational and Experimental Science and Engineering, 11(3). https://doi.org/10.22399/ijcesen.3478

27. Wang G J, Xie C, He K, Stanley H E. 2017. Extreme risk spillover network: application to financial institutions. Quant Finance. 17(9):1417–1433.

28. Billio M, Getmansky M, Lo A W, Pelizzon L. 2012. Econometric measures of connectedness and systemic risk in the finance and insurance sectors. J Financ Econ. 104(3):535–559.

29. Diebold F X, Yilmaz K. 2014. On the network topology of variance decompositions: measuring the connectedness of financial firms. J Econ. 182(1):119–134.

30. Rangu S. (2025). Analyzing the impact of AI-powered call center automation on operational efficiency in healthcare. Journal of Information Systems Engineering and Management, 10(45s), 666–689. https://doi.org/10.55278/jisem.2025.10.45s.666

31. Huang B F, Boutros P C. 2016. The parameter sensitivity of random forests. BMC Bioinf. 17(1):1–13.

32. Friedman J, Posner R. 2011. What caused the financial crisis. Philadelphia: University of Pennsylvania Press.

33. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J C, Müller M. 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinf. 12:77.

34. Hautsch N, Schaumburg J, Schienle M. 2015. Financial network systemic risk contributions. Rev Finance. 19(2):685–738.

35. Rubin R E, Greenspan A, Levitt A, Born B. 1999. Hedge funds, leverage, and the lessons of long-term capital management. Report of The President's Working Group on Financial Markets.

36. Hastie T, Friedman J, Tibshirani R. 2001. The elements of statistical learning. Vol. 1. Springer series in statistics. New York, NY: Springer.

37. Karpman K, Lahiry S, Mukherjee D, Basu S. 2022. Exploring financial networks using quantile regression and granger causality. arXiv. Available from: https://arxiv.org/abs/2207.10705

38. Brownlees C, Nualart E, Sun Y. 2018. Realized networks. J Appl Econ. 33(7):986–1006.

39. Chiodo A J, Owyang M T. 2002. A case study of a currency crisis: the Russian default of 1998. Fed Reserve Bank St. Louis Rev. 84(6):7.

40. Andersen T G, Teräsvirta T. 2009. Realized volatility. In: Handbook of financial time series. Heidelberg: Springer. p. 555–575.

41. Hautsch N. 2012. Econometrics of financial high-frequency data. Heidelberg: Springer.

42. Brennan M J, Jegadeesh N, Swaminathan B. 1993. Investment analysis and the adjustment of stock prices to common information. Rev Financ Stud. 6(4):799–824.

43. Pelger M. 2020. Understanding systematic risk: a high-frequency approach. J Finance. 75(4):2179–2220.

44. Billio M, Getmansky M, Lo A W, Pelizzon L. 2012. Econometric measures of connectedness and systemic risk in the finance and insurance sectors. J Financ Econ. 104(3):535–559.

45. Jha A C. (2025). DWDM optimization: Ciena vs. ADVA for <50 ms global finances. Utilitas Mathematica, 122(2), 227–245. https://utilitasmathematica.com/index.php/Index/article/view/2713

46. Hariharan R. (2025). Zero trust security in multi-tenant cloud environments. Journal of Information Systems Engineering and Management, 10(45s). https://doi.org/10.52783/jisem.v10i45s.8899

47. Jain, R., Sai Santosh Goud Bandari, & Naga Sai Mrunal Vuppala. (2025). Polynomial Regression Techniques in Insurance Claims Forecasting. International Journal of Computational and Experimental Science and Engineering, 11(3). https://doi.org/10.22399/ijcesen.3519

48. Vennamaneni, P. R. (2025). Modernizing monoliths: Transitioning legacy financial systems to microservices. Journal of Information Systems Engineering & Management, 10(45s), 552–574. https://doi.org/10.52783/jisem.v10i45s.8896

49. Singh, V. (2025). Security in Financial Audits. QTanalytics Publication (Books), 21–32. https://doi.org/10.48001/978-81-988770-4-8-3

50. Kumar Tiwari, S. (2023). Security testing automation for digital transformation in the age of cyber threats. International Journal of Applied Engineering & Technology, 5(S5), 135–146. Roman Science Publications.