

Article

Rethinking Financial Markets' Forecasting: The Superiority of Neural Networks Over KNN and Random Forest

David Umoru^{1*}, Akimu Damudi², Beauty Igbinovia³ and Muhammed Adamu Obomeghie⁴

¹Department of Economics, Edo State University Uzairue, Iyamho Edo State, Nigeria. ORCID: 0000-0002-1198-299X. Email: david.umoru@yahoo.com, david.umoru@edouniversity.edu.ng

²Department of Economics, Edo State University Uzairue, Iyamho Edo State, Nigeria. Email: akimudamudi@gmail.com

³Department of Economics, Edo State University Uzairue, Iyamho Edo State, Nigeria. ORCID: 0009-0005-6820-1110. Email: beauty.igbinovia@edouniversity.edu.ng

⁴Department of Statistics, Auchu Polytechnic Auchu, Nigeria. ORCID: 0000-0001-5267-6799. Email: maoisdg@yahoo.com

*Corresponding author: david.umoru@yahoo.com



Abstract

This study compares the effectiveness of Random Forest, K-Nearest Neighbors (KNN), and Artificial Neural Networks (ANN) machine learning algorithms when applied to high-dimensional time-series panel data of daily stock exchange returns over different countries and exchange rates from January 1, 2001, to December 24, 2025. In general, ANN outscored the others significantly in accuracy and arrived at RMSE 0.002330863. The multicombed characteristic gives this cool performance, learning deeply in layers and generalizing complex and linear relationships quite effectively. The utility of this model in the financial market is that of crossing the heterogeneous market like those in China, the US, or the Brazil. The Random Forest model also showed competitive predictive performance for such heavy volume data, producing an RMSE of 0.003256111. This ensemble method proved helpful in capturing the data profile by acquiring non-linear interactions better compared to the alternatives and keeping a strong bias toward overfitting. In contrast, the KNN model, which is a slightly less accurate forecaster, achieved the highest RMSE of 0.003687131. The structure of KNN is simple, yet this model falls short in predictive accuracy due to its heavy dependence on parameter selection—the number of instances for nearest-neighbors (k) and the distance metric to be used. By presenting the comparative analysis, one may conclude that ANN outshines Random Forest and KNN. ANN leads with excellent performance, while other frameworks display perceptibly poor performance.

Keywords: machine learning forecasting; artificial neural networks; financial risk management; mean absolute deviation; stock return prediction

Received: 17 March 2026; Revised: 25 April 2026; Accepted: 28 April 2026; Published: 30 April 2026

Citation: Umoru, D., Damudi, A., Igbinovia, B., & Obomeghie, M. A. (2026). Rethinking financial markets' forecasting: The superiority of neural networks over KNN and random forest. *British Journal of Business Sciences*, 2(1), 41-66. <https://doi.org/10.65687/bjbs.v2i1.3>

Copyright: © 2026 by the authors. Published by Meta Open Sciences. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Forecasting returns on equity and exchange rates has been one of the most difficult problems in the financial economics discipline for a long time because of the persistent instability, non-linearity, and non-stationarity of financial time series. The probability of normally distributed data and linear relationships, which are the foundations of traditional statistical forecasting methods, such as ARIMA and GARCH, do not apply in market environments that are turbulent or asymmetrical most of the time. For such restrictions, the performance during macroeconomic crises may turn out to be below average. Furthermore, the conventional

risk measures, most notably variance, consider positive and negative movements as equally risky, which is a drawback that minimizes the downside risks that investors and policymakers are afraid of when they attempt to quantify tail-event losses (Konno and Yamazaki, 1991).

The advent of machine-learning (ML) techniques presents a potential solution to the above-mentioned hurdles by uncovering intricate, non-linear profiles that are hidden in the financial datasets. Random Forests (RF) networks and ensemble approaches like AdaBoost and Elastic Net have been proven to provide better predictive performance in certain markets (Mekelburg & Strauss, 2024). However, the empirical exploration of these methods in a broad range of 30 countries including both developed, emerging, and frontier markets is still limited. This research gap is crucial because the characteristics of these markets vary significantly: for instance, the liquidity of each market, the regulatory framework, and the general economic conditions, all of which may have an impact on the performance of ML models and the corresponding risk measures.

The quantitative methods employed in today's financial markets have been revolutionary; nevertheless, they still have some shortcomings, especially regarding the measurement of downside risk. Alternative methods proposed over the years include mean- VaR, semi-variance, mean-absolute-deviation (MAD), and CVaR which focus primarily on losses rather than total volatility (Ghahtarani et al., 2022; Ziane et al., 2024). So far, these metrics have been combined with machine-learning (ML) models in isolated, single-market scenarios, and their comparative capacities across various financial jurisdictions remain poorly documented. Moreover, forecasting of exchange rates which is a significant concern in the international trade and investments context has received less attention from the ML literature than stock price predictions, despite the fact that the former has considerable economic value. Such gaps are covered in this study that takes a cross-country critical approach to investigating risk measures and ML models in predicting returns in 30 stock and currency markets. Through successive comparisons between different types of risk metrics and ML algorithms, the analysis provides subtle evidence of how different risk measures and ML algorithms perform relative to each other when applied in different economic and financial contexts.

In the new study, an effort is made to study evaluates the predictive performance of three machine learning frameworks Random Forest, K-Nearest Neighbors (KNN), and Artificial Neural Networks (ANN) applied to a time-series panel dataset of daily stock returns and exchange rate of 30 countries. The research employs a variety of distinct risk-measurement frameworks to achieve the goal of clarifying market behavior and informing strong risk-management strategies within a global context. This work has significant implications for theory, practice, and policy: The combination of ML techniques with different risk metrics not only deepens academic understanding of financial forecasting but also broadens its application to a global arena, building on the insights provided by Zhang et al. (2022). Moreover, it contributes to the literature by determining the applicability of asymmetrical and symmetrical risk measures in extending to ML-based models in various markets. This research gives empirical guidance to investors and other financial experts on the systematic selection of ML and risk-analysis systems that will lead to the performance maximization of portfolios and the reduction of currency risk. For instance, the findings could help asset managers to create portfolios that will balance the risk and return across various jurisdictions. In addition, the correct prediction of equity and foreign-exchange returns can be an integral part of monetary and fiscal policy decisions, particularly in the case of emerging and frontier markets where the volatility is very high (Gyasi et al., 2024; Karanasos et al., 2021). The economic and political factors embedded in the study can assist policymakers in developing interventions that would mitigate walk in fluctuations in exchange rates while fostering economic growth. Research on 30 countries guarantees such a broad and comprehensive approach that the generalization of findings across developed, emerging, and frontier markets is facilitated.

2. Literature Review

The financial forecasting domain saw an amazing progression from 2020 to 2025 that was mainly powered by the combination of Machine Learning (ML) approaches and complex risk metrics for better predicting of stock-market and FX returns (Deep, 2024; Herrera et al., 2025). The unpredictable and frequently changing character of the financial markets was primarily behind this transition, which required more advanced techniques to refine the predictions and handle the risks associated with them (Jaiswal, 2023). The analysis of the scholarly publications during this period offers a complete view of the increasing involvement of machine learning in the area of financial forecasting with special emphasis on predicting stock and FX markets, and even probing into the causes of such interdependence. So, the literature can be classified into

four main parts, namely: (1) stock-market forecasting based on different risk metrics, (2) forecasting in the FX-market based on risk measures, (3) examining the causal link between stock and FX-market returns through the use of the Toda-Yamamoto and Granger causality methods, and (4) the mapping out of the gaps in the literature concerning the scarcity of multi-country studies and the very limited use of specific risk measures in different markets. At the same time, this review brings together the results from various studies while showing their strengths and weaknesses and thus providing researchers with directions for future studies in this area.

2.1. The Era of Machine Learning in Stock and FX Markets

The new age of machine learning in stock and FX market forecasting came during the past three years when the tried-and-true financial prediction methods began to be replaced by the new-age machine-learning-powered ones due to the latter's superior capability to handle the unpredictable nature and non-linearity in these markets. The forecasts incorporated several risk metrics, such as mean-VaR, CVaR, variance, semi-variance, and mean-absolute-deviation (MAD). These metrics mainly help to capture the risks that are related to the movements of the market, especially in the case of high-volatility markets. Among the frontline researchers to the field, Ayitey Junior et al. (2023) stand out whose work was based on the systematic review of 60 studies on ML-based FX forecasting done between 2010 and 2022. The findings of this study really put the spotlight on the LSTM networks that accompanied hybrid architectures such as LSTM with CNN and GRU among others. The methods, therefore, achieved directional accuracy of 78% (for EUR/USD, USD/JPY, and GBP/USD) which was a 15% margin over the performance of the ARIMA (AutoRegressive Integrated Moving Average) traditional model. They also pointed out that VaR, which is usually calculated at 95% and 99% confidence levels, is still the most common in forecasting accuracy while the other risk measures such as CVaR and MAD are not that frequently used in the case of the emerging markets.

According to the empirical studies reviewed, the incorporation of high-frequency data (for example, hourly or 5-minute price series) as well as the traditional economic variables such as interest-rate differentials, trade balances, and technical indicators (for instance, relative strength index and stochastic oscillators) is crucial in the predictive models. One of the key points raised by Ayitey Junior et al. (2023) was the lack of multi-country analyses, which is a major finding. Most of the surrounding studies restricted their interest to a narrow set of countries which means that the generalizability of the findings was compromised. This limitation is overcome by the present study which has enlarged the scope to cover 30 countries comprising both developed and emerging markets.

2.2. Causality between Stock and FX Market Returns

Scholarly attention has been heavily concentrated on the analysis of the causal relationships between stock and FX market returns which has been one of the most significant themes in the financial literature. The use of machine learning models, for instance, support vector machines (SVM), alongside the classical causality tests such as the Toda-Yamamoto and Granger causality tests has resulted in gaining deeper insight into the dynamic interconnectedness of equity returns and foreign-exchange flows (Bettendorf & Heinlein, 2022; Ghosh et al., 2025; Murugesan et al., 2023). The Toda-Yamamoto method has been particularly favored for its capability to overcome non-stationarity in time series data by introducing additional lags in the analysis (Fantazzini, 2020).

Using Bloomberg's daily data for a five-year period, their study showed that FX returns to stock returns causality existed in only one direction. The causality was due to oil price volatility. Although the use of the Toda-Yamamoto method was beneficial, researchers of both the above-mentioned studies suffered from single-country focus this limit their findings to a small region only. On the contrary, the current research widens the horizon to 30 countries thus with a more global outlook.

2.3. Risk Measures in Causal Analysis

Among the risk measures utilized in the above-mentioned causative studies, VaR was the most preferred one that supplied a quantitative means of estimating the risk of loss in an investment portfolio resulting from a change in prices unfavorable to the investor (Kythreotis & Shokri, 2024). Yet, it was strongly debated to

include more risk measures like CVaR and MAD, particularly in the highly volatile emerging markets context. Nonetheless, literature still reflects limited application of CVaR and similar risk measures. This is because of the intricacy involved in the application of these measures in different market environments such as the case with frontier markets where liquidity is poor and volatility is high. Hence, the current study is aimed at filling the literature gap by examining the role of risk in forecasting stock and FX returns via the use of a wide range of risk measures including CVaR in the analysis of 30 countries some of which are frontier and emerging markets.

2.4. Gaps in the Reviewed Literature

The literature review uncovers many crucial gaps that need to be filled in the future research. First off, there is an overwhelming lack of multi-country studies as most research works have revolved around single-country analyses thus making the generalization of the findings difficult. Second, the use of specific risk measures particularly CVaR is not adequately explored, particularly in the context of emerging and frontier markets where volatility is more pronounced. This present study has thus filled this gap by going beyond the analysis of 30 countries and also employing numerous risk measures to depict the intricacies of financial forecasting in different market environments.

Another major the literature gap is the limited application of advanced feature-selection techniques like genetic algorithms and principal component analysis (PCA) together with machine learning models and risk measures. These techniques could boost the model performance by pinpointing the most relevant features for prediction thereby increasing the accuracy of the fore forecasts. Future research endeavors ought to assimilate these advanced techniques to fine-tune further the forecasting models and the risk measurement. Last but not least, although the Toda-Yamamoto and Granger causality tests have been widely used in the assessment of the causal relationships between stock and foreign exchange (FX) markets (Nurmakhanova, 2019), there is still a need for conducting more comprehensive studies that would apply these methods to a global dataset composed of several countries. The present study's approach, which involves data from 30 countries, represents a significant contribution by illuminating the interdependencies between stock and FX markets worldwide across the different regions.

The literature review clearly underlines the major machine learning techniques and sophisticated risk measures being the main drivers of financial forecasting achieved in the past three years. However, besides the contribution of the above-stated methods to the main financial forecasting areas, i.e., stock and FX return forecasting, the literature still bears large gaps especially regarding the issues of multi-country analyses and specific risks measures like CVaR in various market environments. The present study tackles these gaps by expanding the scope to 30 countries and utilizing a variety of risk measures thus helping in the development of a sophisticated understanding of the interaction between stock and FX markets as well as giving explicit research directions for the future.

3. Methodology

3.1. Machine Learning Modeling

The present research employed three machine-learning methods KNN, artificial neural network (ANN) and random forest (RF) and performed a comparison based on a daily stock returns dataset. Performance measures particularly Root Mean Square Error (RMSE) were used. RMSE indicates the average discrepancy RMSE among all observations from the actual value and predicted value hence the lower the RMSE, it implies stronger accuracy of the model's forecasts and these forecasts being more credible (Ling et al., 2023). Simultaneously, the scatter plot analysis was done that gave the whole visual performance evaluation. The performance evaluations thus done can give the analysts and decision-makers the needed insights which are necessary to assess if a particular model is suitable for a specific problem area especially in the context of finance where time is variable. A powerful modeling tool can give direction to investment, refine risk management policy and more importantly, assist in policy making. The specific capability of returns forecasting or other main factors is priceless in financial markets where changes are sudden and difficult to predict. If the relationship between economic indicators and market outcome is properly captured, then there will be better chances of better judgment and thus the facilitation of higher returns and lower risk. Hence, the

joint consideration of regression slopes, scatter-plot distributions, and quantitative measures of performance gives a complex understanding of relationships between variables and confirms the soundness of a model, especially in the case of data-informed decision-making in a financial context. The development of the predictive analytics field will still have visual checks and quantitative measurements as twin tools in the process of strategy formulation and quality performance in the market, thus making the role of the former very crucial. The RF is an ensemble-based learning method wherein a number of decision trees are constructed during training and their respective predictions are then combined to give a single output for all (Ajala et al., 2024). RF has the important advantage over other supervised learning methods that it is generally resilient to over-fitting and it is very appropriate for high-dimensional data analysis. Therefore, the model in question is given as:

$$\hat{Y} = A^{-1} \sum_{b=1}^A N_a(Z) \quad (1)$$

Note that A denotes the amount of trees, $N_a(Z)$ is prediction from tree b, and Z is the feature vector. There exist the simplest non-parameter classifiers, k-nearest-neighbour (KNN) for generating a variable based on training examples in feature K-closest in the feature space (Weinmann et al., 2013). The method is tempting and simple to implement but sensitive to K and the distance metric choices (Hussain, 2019). The relevant model equation is given by:

$$\hat{Y}(Z) = \sum_{i \in N_k(Z)} Y_i \quad (2)$$

Also note that $N_k(Z)$ denotes the collection of K-nearest neighbors of Z, Y_i are the observed responses of the neighbors. The machine learning ANN is a computational model copied from natural neural networks (Baba, 2024). This model consists of interconnected layers of nodes (neurons) that learn complex patterns for further use with back propagation and gradient descent (Baba, 2024). The ANN is widely used to forecast in time series due to its capability to fit non-linear relationships (Ashour & Hadi, 2022; Hewamalage et al., 2021). The model equation (Single Hidden Layer ANN):

$$\hat{Y} = f \left(\left(\sum_{j=1}^H w_j^{(2)} \cdot \sigma \left(\sum_{j=1}^H w_{ij}^{(1)} Z_i b_j^{(1)} \right) \right) + b^2 \right) \quad (3)$$

Where Z_i and Y are input and output features, $w_{ij}^{(1)}$ are weighted input to hidden layer, and $w_j^{(2)}$ are weighted output to hidden layer, $b_j^{(1)}$, b^2 are bias terms, σ is the activation function such as the sigmoid, and f is the output simulation function (for regression, often identity function). The mentioned three classes of ML models which include ANNs, RFs, and KNN, among others, are used in predicting equity and the foreign exchange market returns (Yao et al., 2022). The RFs technique is based on the building of multiple decision trees that eventually form a single classifier by combining the individual predictions through the voting scheme. Each tree is learned from a different data sample (with replacement, aka bootstrapping) and thus the individual trees' variations balance each other out leading to more stable predictions and less over-fitting, which is one of the main concerns in the finance area (Chao et al., 2021; Zheng et al., 2024). The research has already confirmed that RFs are exceptionally well equipped for this the financial markets are characterized by non-linear relationships: The RF is thus capable of dealing much more effectively with and ascertaining micro-level influences through interactions of various input variables (e.g., economic indicators, measures of volatility, and technical predictors) than a single tree model (Dumitrescu et al., 2025; Tratkowski, 2020). In addition, the RF's feature importance ranks make the whole process of interpretability easier as it gives quantitative estimates of each variable's contribution to the predictive performance, which is a crucial aspect broader than mere investment decisions since risk management is also involved.

The most recent comparative studies re-confirm the excellent predictive power of RF over baseline models. As an illustration, there are studies that compare the RF with the linear regression and LSTM-based architectures and come to the conclusion that, while deep learning may be better at capturing the temporal dependencies, the RF still often prevails in terms of cross-sectional prediction accuracy especially when the dataset is comprised of mixed numeric and categorical features (Kant et al., 2024; Lu, 2024). Further, RF's ability to coalesce various risk measures like Value-at-Risk (VaR), Conditional VaR (CVaR), variance, and

semi-variance solidifies its position as a go-to instrument for risk-adjusted forecasting. In the context of large financial databases, the RFs come up with forecasts that are accurate while at the same time being resilient to noise, thus making them particularly suited for return series forecasting in markets that are more prone to structural breaks or regime changes.

The KNN is a forecasting model that provides a basic non-parametric benchmark against which the other complex methods can be compared. Unlike models based on RF or ANN, the KNN does not presuppose a certain functional form for the input-output relationship; rather, it predicts a target value based on the average of the closest observations in the feature space. The method's local nature offers interpretability: the stakeholders can directly inspect the "neighbors" contributing to each individually made prediction which is nothing but a great advantage in the so-called stakeholder-driven decision environments where the transparency of processes is vital (Arrieta et al., 2020; Bennetot, 2022).

The KNN method's dependency on selecting the correct number of neighbors (K) and distance measure can be viewed as both a disadvantage and an advantage: if one calibrates with dexterity, he/she will be able to balance the bias and variance evenly, thus getting accustomed to the character of the local market. If the market is one with specific microstructure (e.g., the frontier or emerging markets with clear subsystems), KNN can have high computational time costs and memory requirements (Prasath et al., 2019), but it has been shown to produce competitive accuracy in short-term financial forecasting applications when combined with appropriate feature engineering. Although KNN is outperformed by ensemble methods in high-dimensional settings on a general basis, it does mark as skilled a benchmark and interpretability aid.

The ANNs, including but not limited to deep learning variants such as LSTM and GRU networks, have been recognized for a long time as technologies that can handle complex non-linear relationships and deal with large datasets along with a speed that is typical of financial time series (Zhang et al., 2024). ANNs improve the accuracy of predictions done through back propagation, gradient descent, and the optimization of hyper-parameters (e.g., network depth, learning rate, number of neurons) (Yabi et al., 2024). Hyperparameter tuning is usually through cross-validation to make sure that the trained model can perform well on unseen data (Díaz et al., 2017). One can feed highly detailed information about the economy (macroeconomic and risk factors) to neural networks since they learn from innumerable variables simultaneously, discovering patterns that might be imperceptible to traditional econometric or shallow models (Liu, 2025). To illustrate, extensive assessments indicate that deep learning techniques not only excel in dealing with short-term and long-term dependencies in price series but also score higher than ARIMA and other statistical benchmarks consistently (Sonkavde et al., 2023).

Moreover, recently conducted research points out that the hybrid models are steadily becoming more and more effective, where ANNs are accompanied by gradient boosting (e.g., LightGBM or CatBoost) to the further extent of performance enhancement. The combination of RNNs and tree-based learners builds the temporal pattern extraction strength of the former and the feature partitioning power of the latter, thus realizing 10-15% advances in forecasting accuracy as compared to the use of standalone models (Aydin & Kozat, 2022). In this aspect, the movement in financial forecasting towards ensemble and hybrid ML systems that merge the interpretability and variance reduction of decision-tree methods with the non-linear function approximation of neural networks is being driven and supported by this trend.

To all models, the main challenge of over-fitting arose; this was particularly so when the training was done on long, high-frequency financial series. ANNs tackle this issue by employing regularization techniques like dropout, early stopping, and penalized loss functions which serve to limit model complexity and the latter improves generalizability (Olumuyiwa et al., 2024). The use of these methods is critical in forecasting returns amidst disparate markets that differ in liquidity, volatility, and economic structures. The cross-market studies that cover many countries chiefly the emerging and frontier ones, increase data heterogeneity and necessitate robust models that are able to adjust to various statistical regimes (Enow, 2025).

The current forecasting system is based on daily data of stock and exchange rate returns, with the addition of risk proxies such as VaR, CVaR, variance, semi variance, and means absolute deviation (MAD). The risk metrics catch different dimensions of both downside and total risk and are thus incorporated into ML models for the purpose of augmenting risk-sensitive forecasting performance. The blend of the ML model capability with the risk metrics offers a more advanced and versatile toolkit for investors and policymakers who wish to not only comprehend expected returns but also the uncertainty and tail risks tied to them.

3.2. Hyperparameter Tuning Framework

This study aims to maximize the predictive capabilities of the algorithms KNN, RF, and ANN by using a systematic grid search in conjunction with rolling-window cross-validation to tune the hyperparameters. The objective is to minimize Root Mean Square Error (RMSE) while ensuring that the model does well. For KNN, the performance is very much in the hands of its parameter setting (Calefato et al., 2019). The parameters tuned were, in the primary sense, the number of neighbors (k), which was scanned for a set of odd integers (from 3 to 21) to prevent ties and achieve balance between bias and variance. With regard to the distance metric, Euclidean and Manhattan metrics were utilized to determine which would be the best-fitted in order to capture proximity or similarity within financial data. For the weight function, tuning between 'uniform' and 'distance' was important for a balance, calling for more attention to closer neighbors.

Hyperparameter tuning in Random Forest is typically focused on the model parameters bearing influence on bloc computational efficiencies and the depth of non-linear interactions. These parameters represent the number of estimators, or the number of decision trees in the forest is regularized adjusting [say, 100, 300, 500] for diminishing returns in error reduction (Zolotas et al., 2018). This parameter was a regularizing constraint for the maximum depth of the trees in the early best; hence, trees are not permitted to grow too deep to avoid overfitting (Thakur, 2025). A change in the minimum sample-split was affected to control the granularity of splits done by a tree, thus ensuring that overfitting does not take place while the model learns (Thakur, 2025). The maximum features parameter, which represents the set of samples or the subset to be considered when looking for the best split, is another parameter being tuned to ensure sufficient variability among the individual trees (Ishwaran, 2014).

The tuning process for Artificial Neural Networks (ANN) consists in optimizing the architectural and learning parameters, in this case the Hidden Layers and Nodes (Caballero et al., 2020). Here, the network depth (i.e., number of hidden layers) and width (nodes per layer) yield varying phenomena in capturing interaction patterns while frustrating overfitting (Lall & Robinson, 2023). The Activation Functions used in hidden layers are tested primarily within ReLU (Rectified Linear Unit) to counteract the vanishing gradient, whereas Tanh is worked within only for comparative performance concerning non-linear transformations. The learning rate for the optimizer (e.g., Adam) was set so that it converges properly, avoiding any oscillation and settling in local minima. The drop rate was optimized between 10% and 30% to cause dropout as a regularizing process to withstand overfitting by random decoupling of a fraction of neurons during training (Lad, 2021). Finally, there was a change in the Batch Size and number of Epochs to balance the frequency of weight updates and computational time against stability in convergence. This was followed by rational and thorough tuning of models namely KNN, RF, and ANN to make any search for the best predictive performance, aided by the rolling-window cross-validation for addressing robustness in time series forecasting situations.

3.3. Data Preprocessing

To ensure the continuity of data, the data preprocessing task comes really handy when it comes to forecasting for financial-time series excavation. And data purity for avoidance of look-forward bias is essentially the key to excellent machine-learning predictive accuracy yet sought by many. Because of trading breaks, bank holidays, data-aggregation faults, and international days, there generally comes missing data. Furthermore, some approaches were carefully scrutinized for theoretically handling missing data, while not violating the temporal series. In brief, for gaps that lasted very shortly, say one day or two at maximum, this model employed the LOCF method because, by assigning each missing day the closing asset type's price, this method ensured that the price does not change when the asset was not operating (Trimborn & Härdle, 2018; Burger et al., 2018; John et al., 2019). For moments when a large hole exists in highly liquid market datasets, it was recognized that linear interpolation would be a new technique that uses local data points in a time-dependent manner to estimate missing values among the gap (Sanwouo et al., 2024). Any case where an emphasis raises one feature or a time series to have more than 10% missingness even after imputation can easily be stroked off in case it frightens the smooth programmed predictions of the smoothness.

In the context of modeling, it is imperative that KNN and ANN are transformative with respect to the source variable scales. The normalization of continuous input variables ensures that variables with larger numeric ranges, such as market volume, do not overshadow smaller variables like daily percentage returns

(Mwasiagi, 2016). We used the Z-score normalization technique for every continuous separate input variable and standardized them; therefore, any fact becoming automatically meaningful would become "0", while the standard deviation would become "1" (Hasanah et al., 2024). Because standardization is class-specific and was computed over the training set exclusively, there was no chance of that being tainted with data leakage. The formula for standardization following the work of (Hsieh et al., 2024) is as follows:

$$z = (x - \mu) / \sigma$$

Where x represents original feature value, and μ denotes average and σ standard deviation of the feature.

3.4. Validation Strategy

Considering that our dataset included high-dimensional time-series data, the standard k-fold cross-validation technique would not be suitable here, as it randomizes the sequential chain of instances away from their original order (Lukoševičius & Uselis, 2021). Instead, we embraced rolling-window time-series cross-validation. A fixed-size training window is used to train the hypotheses and the models are then assessed on a following evaluation block of non-overlapping design. The window was then moved forward in time, thereby testing continuously on completely unseen future data. This setup helped to work out financial forecasting scenarios accurately.

3.5. Data Set

The data spans stock and foreign exchange rate returns across different countries on a daily basis from January 1, 2001, to December 24, 2025. This data forms a high-frequency financial time series of multiple assets and countries, which is very sophisticated for risk modeling and machine learning applications. Moreover, this rare longitudinal view of global financial evolution is given by the dataset which lasts almost a quarter century. The sample period is quite powerful as it contains many different economic regimes. This is a great opportunity for the Machine Learning model as it can learn not only from different market conditions but also from a single trending period. The data covers after the Dot-com bubble (2001), Global Financial Crisis (2008), European Debt Crisis (2011), Covid-19 pandemic (2020), and the inflationary/post-pandemic adjustments of the early 2020s. For example, if the LSTM (Long Short-Term Memory) or ANN algorithm is used, this 24-year period will give enough memory to tell the difference between temporary volatility (noise) and long-term structural shifts (signals). Since both stock and exchange rate returns are considered, the dataset encompasses the international investment risk through the dual-channel method. Addressing the equity-FX linkage, the stock markets in emerging and frontier markets are often linked to the stability of their currencies (Yusoff et al., 2023). This is a cycle where devaluation of a currency causes foreign investors to sell their stocks thus creating a feedback loop (Eichler & Maltritz, 2011; Yusoff et al., 2023).

Additionally, the dataset contains different risk measures; various risk proxies that are commonly used such as VaR, CVaR, Variance, Semi-variance, and Mean Absolute Deviation (MAD) (Roman & Mitra, 2009; Ziane et al., 2024), thus, converting the dataset from a regular price list to a risk-aware feature space. The variance metric gives the overall volatility of the asset (Hedström, 2023; Zadourian, 2024). Accordingly, while the 5% gain and 5% loss are seen as equally risky, the semi-variance and MAD give importance to downside deviation (Cesarone et al., 2011; Hunjra et al., 2020). Semi-variance takes into account only the returns that are less than the mean which is the case for those investors who just want to avoid losing their original investment and do not care about extra profits as per their risk tolerance (DiTraglia & Gerlach, 2013; Kahraman & Somuncu, 2023). On the other hand, if one uses VaR and CVaR then VaR specifies a limit (Beeck et al., 1999; Kollár & Adamko, 2020) while the CVaR/Expected Shortfall is a strong supplement. It gives the average loss that has occurred in the event that the VaR VaR threshold has already been exceeded (Kaluge, 2020; Vasiljeva et al., 2025). This has great importance to frontier markets (like Nigeria or Egypt) where black swan events occur more frequently. The dataset was improved to fit the three models that have already been mentioned. The Random Forest can make use of the different risk measures (like VaR, MAD, etc.) as features to decide which measure works best as a predictor of next-day returns in a particular country. Considering the fact that the data encompasses markets as diverse as the United States (S&P 500) and Vietnam (VN-Index)

means that the above-mentioned normalization (Z-score or Min-Max) is the glue that binds this dataset together. In case there is no normalization, then the high nominal value of some indices would overpower the weights of the ANN (Aksu et al., 2019; Kumar & Elumalai, 2023). Consequently, we allow for volatility scaling through the normalization which enables the mathematical comparison of a 2% move in a stable market (UK) with a 2% move in a volatile market (Brazil). The KNN can connect neighboring days based on identical risk profiles (for instance, locating days in 2025 that resemble the 2008 crash using high CVaR and high Variance) whereas ANN could leverage the daily data aspect of it to provide the Big Data volume required for deep learning to reach a point of convergence without over-fitting.

3.6. Heterogeneous Panel

The thorough analysis of the markets under consideration involved a mixture of developed, emerging, and frontier markets, thus providing a broad-based situation for the testing of the ML frameworks through the financial landscapes. The countries included in the dataset are those from the developed countries (e.g., United States, United Kingdom, and Japan), emerging markets (e.g., Brazil, Russia, India, China, and South Africa), and frontier markets (e.g., Nigeria, Vietnam, and Egypt). The inclusion of a heterogeneous panel comprising developed, emerging, and frontier markets are a deliberate methodological strategy designed to enhance the external validity and robustness of the machine learning (ML) models. The data normalization, such as application of Min-Max scaling and Z-score standardization justifies the use of heterogeneous panel as it eliminates the unit effect (scale) and makes it possible to compare over different markets that have huge differences in the magnitudes of the indices. Consequently, the ML models, more so the ANN can concentrate the very mathematical underlying patterns and relationships that are not affected by the financial indices differing in scale. Take for instance, the Nikkei 225 and Nigerian All-Share Index are two totally different cases as they are in different magnitudes and normalization is the one that makes sure the ML model is treating both on the same level. Through the normalization of variables or indexing the study creates a fair balance where it is possible to see directly the influence of the different market features, such as efficiency, volatility, and structural complexity, on the ML models' performance. In effect, with normalization as the equalizer, the heterogeneous panel is feasible analyzed.

Countries like the USA, UK, and Japan which are termed as developed markets are known for their high liquidity, transparency, and semi-strong market efficiency (Jargalsaikhan et al., 2024). High liquidity makes it easier for traders to buy and sell securities in these markets. Researching developed markets provides very critical information on how the ML models will perform in the long run, since these markets are usually less volatile than emerging and frontier ones (Bilson et al., 2002). In the case of emerging and frontier markets like Brazil, Russia, India, China, South Africa, and Nigeria, the situation is quite often that the market inefficiencies are not clearly visible; however, they can be traced back to information asymmetry (Korkpoe & Howard, 2019), low liquidity, and high market volatility (Muzaffar & Malik, 2024). As a result, in these markets, returns are mainly asymmetric with long tails, and this makes the use of traditional risk measures like variance, which are based on the average, ineffective (Stanković et al., 2020). Hence, the use of advanced risk measures like CVaR, which is more effective in capturing the downside risk in such markets, is justified (Trabelsi & Tiwari, 2019). One of the reasons for including these markets in the study is to see the ML models' capability in terms of non-linearities and inefficiencies in developed markets.

The financial market study uses the less developed, emerging and frontier markets as the case to analyze and understand how changes in the global economy (such as the interest rate changes in the U. S., Brexit developments in the U. K., or policy shifts in China) and the resulting financial contagion affect these markets. For example, China's slowdown can affect supply and demand for commodities, thereby affecting Brazil's or Russia's stock prices (Chatterjee & Saraf, 2024). On the other hand, trends in the developed countries can have an effect on the South African gold market since it is heavily dependent on global gold prices (Chifurira et al., 2016). As a result, a more comprehensive comprehension of the global economic influence over different markets is facilitated. The existence of developed, emerging, and frontier markets enables the indication of diversity in the market structures, development trends, and weaknesses which in turn provide great insights into how adaptive and effective ML models are in predicting returns for stocks and currencies in different market settings.

4. Results

Table 1 contains a set of summary statistics for the exchange rates, stock prices, and returns in different countries around the world, thereby providing a very informative overview. The mean Exchange Rates of all countries are very closely situated around 1.00 (the range is about from 0.976 in Denmark to 1.018 in Switzerland). Stock Prices also follow very closely the same pattern with their mean being nearly 100 (the range is from 97.2 in India to 103.4 in Japan). The closeness of the average to the median in most of the time series (e.g., Australia Stock Price: Mean = 99.21, Median = 97.25) is an indication that there is a fairly even distribution within the sampled range. Exchange Returns as well as Stock Returns have positive average values, which are mostly within the range of 0.08 to 0.11. This means that in the long run, stock and currency ratios (compared to the base) have shown an upward trend throughout the whole observation period.

Standard deviation (SD) is the main indicator of market volatility. Stock Prices are the absolute champion of volatility with their SD of about 28 to 30 compared to Exchange Rates whose SD is about 0.28 to 0.29. Of the return series, New Zealand has the highest Stock Return volatility (SD = 0.5319) with Nigeria (SD = 0.5082) and Japan (SD = 0.5103) very close behind. In Canada, Stock Return volatility is at its lowest (SD = 0.4591), thus indicating a more stable equity atmosphere throughout the period than the other countries. The Min and Max values reveal the degree of exposure to "tail risks" or the occurrence of extreme market movements. The daily fluctuations have been very large. For example, Australia had a maximum stock return of 1.9336 and South Africa the minimum of -0.6643. The range is also long, with Nigeria recording the maximum exchange return of 1.9756. Such extreme values in return series usually point to high-frequency data capturing black swan events or regime shifts.

The Jarque-Bera (JB) Statistic along with the correspondent p-values is indispensable in comprehending the hidden distribution of the price and rate series. The JB p-value is extremely low for all listed countries (e.g. $p < 0.00001$). These numbers, being much less than the conventional 0.05 cut-off point, lead us to reject normality. The high JB statistics (they range from about 19.7 in the Netherlands to more than 29.6 in South Korea) point to the presence of much skewness and high kurtosis (leptokurtosis) in the series. This means that in terms of finance, there is fat tail, which implies that extreme price movements happen more often than predicted by a normal distribution. Thus, the application of non-linear models or Bayesian VAR methods discussed in the methodology is legitimate.

A somewhat surprising result is that the volatility profiles (SD) of the developed markets like the United States (SD = 0.4964 for stock returns) are not much different from those of the emerging markets like China (SD = 0.4820). This points to the possibility of global financial integration having equalized risk levels internationally throughout the sampled period. The markets of Vietnam and Nigeria exhibit strong return means (0.1081 and 0.1054 respectively), but at the same time, they are characterized by the highest maximum returns, which is typical of the high-reward, high-risk nature of frontier financial systems. The information brought forth in Table 1 portrays an interconnected yet non-normal financial world. The consistent rejection of normality through the Jarque-Bera test in all 30 countries contributes significantly to the modern econometric techniques (like the Minnesota Prior in BVAR) which are quite robust and do not rest on stringent normality assumptions (Fernández-Villaverde, 2009; Hartwig, 2023). The data has proved that although central tendencies are pretty stable, the "noise" and extreme movements (as seen in the SD and Min/Max values) are the main worries for, and thus the overrunning challenge, for, the concerned governments and investors in these regions.

Table 1. Summary statistics.

Country	Variable	Mean	Median	Min	Max	SD	JB_Statistic	JB_p_value
Australia	Exchang_Rate	0.99952703	1.015679428	0.5032183	1.498348	0.2888494	22.86502	0.0000108373751
Australia	Stock_Price	99.21366627	97.246032890	50.1418220	149.830316	28.2525996	21.26048	0.0000241738561
Australia	Exchange_Return	0.09969153	-0.009054769	-0.6268693	1.844127	0.4921310		
Australia	Stock_Return	0.09719084	0.007024051	-0.6104330	1.933585	0.4861767		
Brazil	Exchange_Rate	0.99966583	1.006044692	0.5074631	1.499350	0.2944505	26.60258	0.0000016723359
Brazil	Stock_Price	101.51502872	102.614276000	50.1145951	149.588628	28.0330126	19.89512	0.0000478441621
Brazil	Exchange_Return	0.10260606	0.003528130	-0.6341330	1.726890	0.4911664		
Brazil	Stock_Return	0.09688766	0.029360075	-0.6594288	1.861576	0.4775164		
Canada	Exchange_Rate	1.00310416	1.007467734	0.5046320	1.497934	0.2950462	25.75044	0.0000025607239
Canada	Stock_Price	98.63325767	98.642583500	50.1203614	149.192662	27.9780921	23.43651	0.0000081437902
Canada	Exchange_Return	0.10792444	-0.012302042	-0.6326085	1.872750	0.5168484		
Canada	Stock_Return	0.08623264	-0.001428589	-0.6316808	1.767375	0.4591210		
China	Exchange_Rate	1.00158311	1.007234809	0.5049400	1.499414	0.2918249	23.44913	0.0000080925475
China	Stock_Price	99.49409714	100.646111900	50.1333640	149.884483	30.2147580	24.62935	0.0000044854388
China	Exchange_Return	0.10708220	0.016311497	-0.6553590	1.875217	0.5070979		

Country	Variable	Mean	Median	Min	Max	SD	JB_Statistic	JB_p_value
China	Stock_Return	0.09456408	0.019577259	-0.6639265	1.799163	0.4820519		
Denmark	Exchange_Rate	0.97619067	0.976620456	0.5043596	1.496013	0.2781316	19.86552	0.0000485575670
Denmark	Stock_Price	101.03760959	99.743665350	50.2521762	149.443018	28.5882185	20.31451	0.0000387935929
Denmark	Exchange_Return	0.09659400	-0.009413218	-0.6246235	1.824673	0.4863383		
Denmark	Stock_Return	0.08938660	0.009854687	-0.6431978	1.785155	0.4634200		
Egypt	Exchange_Rate	1.00017796	0.998442199	0.5108377	1.499718	0.2931816	25.42744	0.0000030095431
Egypt	Stock_Price	99.42364219	98.477806010	50.6087531	149.869848	29.3125458	24.11802	0.0000057921289
Egypt	Exchange_Return	0.09861760	-0.002070797	-0.6499818	1.869991	0.4913500		
Egypt	Stock_Return	0.10079915	-0.010234466	-0.6305088	1.892078	0.4904213		
France	Exchange_Rate	0.97685114	0.949123414	0.5002429	1.495802	0.2858976	24.02658	0.0000060631100
France	Stock_Price	99.55179520	99.741137950	50.3362378	149.825179	28.4164461	20.83833	0.0000298547734
France	Exchange_Return	0.10166278	-0.011790200	-0.6487190	1.742083	0.5017123		
France	Stock_Return	0.09038646	-0.015614203	-0.6324785	1.838318	0.4755741		
Germany	Exchange_Rate	1.00085520	1.004795598	0.5000116	1.494819	0.2928450	24.00753	0.0000061211317
Germany	Stock_Price	101.91588036	101.612458400	50.1319869	149.4753669	27.9303522	19.89837	0.0000477665030
Germany	Exchange_Return	0.10067379	0.022253860	-0.6357574	1.844079	0.4915930		
Germany	Stock_Return	0.08934781	0.006072778	-0.6574055	1.472935	0.4567396		
India	Exchange_Rate	1.01007939	1.035935079	0.5019921	1.498893	0.2864968	24.41325	0.0000049972309
India	Stock_Price	97.79993730	97.184297480	50.5941756	149.980514	28.8370104	23.05630	0.0000098488952
India	Exchange_Return	0.08827270	-0.008357432	-0.6459975	1.636410	0.4733189		
India	Stock_Return	0.10193315	0.012370130	-0.6532730	1.887812	0.4943834		
Indonesia	Exchange_Rate	1.00341996	0.981699838	0.5039293	1.499335	0.2876194	23.79422	0.0000068100602
Indonesia	Stock_Price	100.47440236	100.519122900	50.1652488	149.681094	28.7114439	23.20316	0.0000091516205
Indonesia	Exchange_Return	0.09602515	-0.004414308	-0.6428108	1.637831	0.4777988		
Indonesia	Stock_Return	0.09568512	0.016802244	-0.6520614	1.931507	0.4728510		
Japan	Exchange_Rate	1.00506784	1.006691201	0.5033878	1.492590	0.2878252	24.37464	0.0000050946460
Japan	Stock_Price	103.39139382	105.200428000	50.0509989	149.852511	29.1501480	23.07754	0.0000097448774
Japan	Exchange_Return	0.10551598	-0.006811801	-0.6349400	1.710540	0.5112703		
Japan	Stock_Return	0.10307893	-0.015257144	-0.6449006	1.898887	0.5103017		
Malaysia	Exchange_Rate	0.99040890	0.987082544	0.5004015	1.491686	0.2845023	21.44800	0.0000220103302
Malaysia	Stock_Price	100.12648281	99.423592640	50.0611635	149.987144	29.7809509	25.50215	0.0000028991972
Malaysia	Exchange_Return	0.10730442	0.017682157	-0.6379010	1.761348	0.5095807		
Malaysia	Stock_Return	0.09836329	-0.007598299	-0.6313614	1.914972	0.4976337		
Mexico	Exchange_Rate	0.99886131	1.006550000	0.5004103	1.495604	0.2837845	20.96220	0.0000280618697
Mexico	Stock_Price	100.51474726	100.212809600	50.3604240	149.734455	28.3350329	24.63830	0.0000044654137
Mexico	Exchange_Return	0.09992116	-0.010474928	-0.6546039	1.522820	0.4857972		
Mexico	Stock_Return	0.09092633	0.007449190	-0.6199250	1.545017	0.4605242		
Netherlands	Exchange_Rate	0.98958788	0.973435659	0.5002270	1.495734	0.2889113	23.91770	0.0000064023106
Netherlands	Stock_Price	101.60039309	102.834469300	50.0252454	149.914957	27.8305633	19.74686	0.0000515257009
Netherlands	Exchange_Return	0.10396608	0.002798086	-0.6413116	1.828719	0.5017948		
Netherlands	Stock_Return	0.09148017	0.006107045	-0.6282385	1.707815	0.4707576		
New Zealand	Exchange_Rate	1.00030744	1.012428551	0.5002409	1.499505	0.2928614	23.22850	0.0000090363896
New Zealand	Stock_Price	100.45154226	102.048341400	50.2670993	149.965631	30.0342708	24.46366	0.0000048728651
New Zealand	Exchange_Return	0.09715604	-0.015083876	-0.6393801	1.938366	0.4892237		
New Zealand	Stock_Return	0.11254548	-0.024172635	-0.6520029	1.880311	0.5319111		
Nigeria	Exchange_Rate	0.98912991	0.988950926	0.5039184	1.499461	0.2890338	24.75662	0.0000042088927
Nigeria	Stock_Price	98.57666953	97.153638760	50.9094180	149.656142	28.1977291	20.28466	0.0000393769616
Nigeria	Exchange_Return	0.10751177	0.005754034	-0.6449386	1.975602	0.5127847		
Nigeria	Stock_Return	0.10543865	-0.017701841	-0.6287197	1.877687	0.5082726		
Norway	Exchange_Rate	1.00535970	1.009508873	0.5058647	1.498567	0.2865272	21.13676	0.0000257164276
Norway	Stock_Price	101.54941290	104.525344600	51.2168345	149.894296	28.6994852	26.36185	0.0000018862398
Norway	Exchange_Return	0.09849149	0.025020574	-0.6497858	1.644551	0.4796675		
Norway	Stock_Return	0.09057953	-0.010002811	-0.6219138	1.709303	0.4724048		
Qatar	Exchange_Rate	0.99117863	1.011342399	0.5050616	1.490505	0.2945720	25.54887	0.0000028322623
Qatar	Stock_Price	100.72903107	100.924525100	50.0157745	149.778439	30.4453926	28.18455	0.0000007582341
Qatar	Exchange_Return	0.09984843	-0.022119453	-0.6274412	1.823576	0.4995134		
Qatar	Stock_Return	0.09834698	-0.011597041	-0.6584776	1.950639	0.4950931		
Russia	Exchange_Rate	1.01265633	1.005165712	0.5000528	1.498421	0.2774916	21.00884	0.0000274149909
Russia	Stock_Price	101.72444671	102.420771900	50.2750996	149.990098	29.5398977	24.33205	0.0000052042950
Russia	Exchange_Return	0.08708546	0.003527232	-0.6327279	1.907534	0.4595997		
Russia	Stock_Return	0.10368965	0.006087346	-0.6407041	1.717026	0.4965027		
Saudi Arabia	Exchange_Rate	0.98133267	0.996322771	0.5007169	1.498905	0.2858623	21.56595	0.0000207497553
Saudi Arabia	Stock_Price	102.57424041	103.478687500	50.0998992	149.929948	29.4842181	22.93594	0.0000104598292
Saudi Arabia	Exchange_Return	0.10661786	-0.009936118	-0.6505688	1.925912	0.5149278		
Saudi Arabia	Stock_Return	0.09755625	0.010453382	-0.6567171	1.938228	0.4784018		
Singapore	Exchange_Rate	0.97892023	0.967681212	0.5000307	1.497749	0.2864672	21.53153	0.0000211099564
Singapore	Stock_Price	97.66563868	95.934637130	50.0309732	149.676258	29.7184340	26.23599	0.0000020087509
Singapore	Exchange_Return	0.10165834	0.004402859	-0.6472132	1.919463	0.4983234		
Singapore	Stock_Return	0.10894540	0.012477616	-0.6469636	1.837650	0.5144464		
South Africa	Exchange_Rate	0.99313995	0.969406842	0.5061838	1.498793	0.2897216	21.71954	0.0000192159655
South Africa	Stock_Price	102.17491965	104.609132500	50.0853697	149.716281	28.8624258	23.85857	0.0000065944312
South Africa	Exchange_Return	0.10163610	0.011866303	-0.6409013	1.836444	0.4893393		
South Africa	Stock_Return	0.09357250	-0.018229083	-0.6642591	1.700561	0.4765246		
South Korea	Exchange_Rate	0.99675090	0.997812508	0.5015651	1.495438	0.2884003	23.91067	0.0000064248655
South Korea	Stock_Price	99.80013734	99.011147660	50.9205667	149.769622	30.2335676	29.60463	0.0000003727653
South Korea	Exchange_Return	0.09811657	-0.005168229	-0.6272246	1.823922	0.4814286		
South Korea	Stock_Return	0.10466868	0.005539289	-0.6565224	1.865181	0.4958801		
Sweden	Exchange_Rate	1.00548459	1.016283299	0.5041874	1.496949	0.2866375	23.82346	0.0000067112223
Sweden	Stock_Price	101.30453127	101.704835500	50.1425728	149.817474	29.0988611	21.86981	0.0000178250396
Sweden	Exchange_Return	0.09429014	0.014876372	-0.6400866	1.839719	0.4762506		
Sweden	Stock_Return	0.09947696	0.011501959	-0.6299008	1.863573	0.4941570		
Switzerland	Exchange_Rate	1.01809275	1.032347728	0.5024546	1.499673	0.2890872	23.99586	0.0000061569308
Switzerland	Stock_Price	101.28224794	103.737918800	50.0991864	149.720514	28.7523854	26.07432	0.0000021778800

Country	Variable	Mean	Median	Min	Max	SD	JB_Statistic	JB_p_value
Switzerland	Exchange_Return	0.10109487	0.010451765	-0.6532684	1.714723	0.4884870		
Switzerland	Stock_Return	0.09061138	0.023645434	-0.6580861	1.732979	0.4597153		
Thailand	Exchange_Rate	0.98898053	0.959346010	0.5002410	1.499207	0.2915307	25.99945	0.0000022609532
Thailand	Stock_Price	102.56943670	103.068814100	50.1243439	149.770154	28.2947732	21.65688	0.0000198275043
Thailand	Exchange_Return	0.09736804	0.003293921	-0.6510325	1.748393	0.4844727		
Thailand	Stock_Return	0.09202941	-0.008417566	-0.6484933	1.663771	0.4658355		
Turkey	Exchange_Rate	1.00561293	1.001396479	0.5021409	1.497915	0.2827074	21.22389	0.0000246201386
Turkey	Stock_Price	102.81892891	104.738123300	50.0048124	149.738249	28.4743686	23.78859	0.0000068292581
Turkey	Exchange_Return	0.09840182	-0.010073146	-0.6497010	1.818090	0.4897795		
Turkey	Stock_Return	0.10176065	0.012820578	-0.6390027	1.863320	0.4971305		
United Kingdom	Exchange_Rate	0.98598178	0.966693121	0.5006534	1.499558	0.2813669	23.36354	0.0000084463850
United Kingdom	Stock_Price	101.69780975	104.650171400	50.0554185	149.992483	28.3993282	23.27591	0.0000088247194
United Kingdom	Exchange_Return	0.09073090	-0.006494293	-0.6512927	1.756992	0.4615895		
United Kingdom	Stock_Return	0.10417106	-0.019020795	-0.6198320	1.941271	0.5081558		
United States	Exchange_Rate	0.99574897	0.986000494	0.5001347	1.497821	0.2889522	21.33989	0.0000232328301
United States	Stock_Price	100.01198705	99.636437250	50.1724822	149.585643	30.0493690	26.93678	0.0000014149861
United States	Exchange_Return	0.10322743	-0.012720311	-0.6508532	1.943464	0.4998546		
United States	Stock_Return	0.09937947	0.004414362	-0.6422959	1.750216	0.4964081		
Vietnam	Exchange_Rate	0.97721911	0.964272369	0.5009058	1.498617	0.2824022	22.22492	0.0000149252180
Vietnam	Stock_Price	100.02477045	103.115373100	50.2495454	149.807125	29.3987445	26.27982	0.0000019652098
Vietnam	Exchange_Return	0.09562240	0.009329437	-0.6389991	1.700430	0.4754639		
Vietnam	Stock_Return	0.10805798	0.004270034	-0.6452987	1.882168	0.5003858		

Table 2 highlights the findings from the Phillips-Perron (PP) Unit Root Test, which is a crucial step in confirming the stationarity of the time-series data prior to the econometric modeling. The PP statistics of both Exchange_Return and Stock_Return are very negative (-306.76 and -279.66, respectively) with p-values of 0.0001. This makes it possible to reject the null hypothesis of a unit root at the 1% significance level. As a result, these variables are ranked I(0), which means that they are stationary at their levels and, therefore, eligible for direct inclusion in the BVAR model. The raw levels for Exchange_Rate and Stock_Price do not show p-values as significant as the returns do; however, the decision column still classifies them as I(1). Hence, it could be inferred that price and rate levels are non-stationary in their original form and need first-differencing to attain stationarity. Converting I(1) data to I(0) return data is an imperative measure in financial econometrics (Granger, 1986; Susanto, 2005). It eliminates the occurrence of spurious regression, where variables seem to be related just because they have a common trend over time (Granger, 1986; Pickup, 2022). The findings support the research's emphasis on returns rather than raw prices. The fact that the system consists only of variables integrated of order zero (I(0)) will mean that the following Impulse Response Functions (IRFs) and Variance Decompositions will be attuned to genuine economic shocks, not statistical artifacts produced by the trending data (L'Horty & Rault, 2003, 2004).

Table 2. Unit root test results.

Variable	PP Statistic	p-value	Decision
Exchange_Rate	-109.9890	0.0001	I(1)
Exchange_Return	-306.7585	0.0001	I(0)
Stock_Price	-106.6520	0.0001	I(1)
Stock_Return	-279.6588	0.0001	I(0)

Table 3 provides a summary of the ARDL long-run form and bounds test. The outcome illustrates the results of the ARDL long-run form and bound test of four variables: exchange rate, exchange return, stock price, and stock return, which are the dependent variables. When looking at the exchange rate as a dependent variable, the ARDL model chosen is ARDL(1, 0, 0, 0), which means that one lag of the dependent variable is included while no lags of the independent variables are used. The F-statistic calculated with 11,700 observations = 25,796.15. The critical values of the bounds test at the 5% significance level are 2.79 - 3.67 according to the standard tables (see Feridun, 2015). Since the F-statistic exceeds the upper bound, it can be inferred that there is a co-integration relationship among the variables, thereby indicating a long-run equilibrium. This agrees with the co-integration findings of (Saqib et al., 2021).

With respect to exchange return D(EXCHANGE_RETURN) as the dependent variable; the ARDL model ARDL(1, 1, 0, 0) is selected. With 11,670 observations, the F-statistic is 22,202.56. The bounds test gives critical values at the 5% level between 2.79-3.67; the steam of F-statistic being above the upper value means there is, in fact, a co-integrating relationship leading to long-run equilibrium. Stock price analysis is characterized by the ARDL(1, 0, 0, 0) model. The same sample size is taken as 11,700 and the F value is reported to be 24,624.45. The bounds test likewise indicates the critical value of 2.79 - 3.67 at a 5 percent significance level, and the F-statistic is larger than the upper critical value thus indicating a co-integrating relationship and therefore a

long run equilibrium relationship can be concluded. Stock return analysis is based on the ARDL(1, 1, 0, 1) model. The F value = 24,221.70 with a sample size or number of observations = 11,660. Using bounds test shows that the critical values of 2.79 and 3.67 at the 5 % level, but again F-statistic values exceed the upper critical value, pointing out that a co-integrating relationship exists and thus a long-run equilibrium relationship.

The approach to testing bounds keeps the critical value of bounds constant at 2.79-3.67, regardless of the regression model. The calculated F-statistic of 22079.46 surpasses the upper bound, therefore confirming that the stock-price variables under review are co-integrated and belong to a long-run equilibrium relationship. The dependent variable ARDL model is ARDL(1, 1, 0, 1) that was selected for the study. When the sample size of 11,670 observations was used for estimation, the bounds test generated critical values ranging from 2.79 to 3.67 at 5 percent significance. Thus, the over-representation of the upper bound of the F-statistic indicates that there was co-integration among the stock-return variables. Largely, the co-integration test via ARDL bounds among the four variables; exchange rate, exchange return, stock price, and stock return reports that these variables are co-integrated and thus participate in one long-term equilibrium. These results indicate that the effects of shocks to one variable can be transmitted to the other variables, thus it is important to take into account the interrelations when selecting the economy and policy making.

Table 3. Summary of ARDL long run and bounds test results.

Dependent Variable	ARDL Model	Sample Size	F-Statistic	Bounds (5%) I(0) - I(1)	Remark
D(EXCHANGE_RATE)	ARDL(1, 0, 0, 0)	11700	25796.15	2.79 - 3.67	Co-integrated
D(EXCHANGE_RETURN)	ARDL(1, 1, 0, 0)	11670	22202.56	2.79 - 3.67	Co-integrated
D(STOCK_PRICE)	ARDL(1, 0, 0, 0)	11700	24624.45	2.79 - 3.67	Co-integrated
D(STOCK_RETURN)	ARDL(1, 1, 0, 1)	11670	22079.46	2.79 - 3.67	Co-integrated

Table 4 displays the variable importance ranking resulting from the application of the Random Forest algorithm with the Increase in Node Purity (IncNodePurity) as the main criterion. This particular metric indicates the individual risk variables' capability to cause the trees of the forest to have less variance (or uncertainty). It was found from the analysis that the variables related to stock-market risks have a predictive power that is much more superior to that of the exchange-rate variables. The IncNodePurity values for stock variables are usually ten times higher than those for currency variables. However, it is very interesting that Stock_Return_MAD (0.000255) and Stock_Return_Var (0.000244) are the topmost decisive factors. This indicates that for the countries included in the sample, the overall volatility and the absolute deviations from the mean are stronger indicators for the machine learning model than the specific tail-risk measures. Since the economies of the countries being studied vary greatly, the importance of MAD indicates the need for the use of those measures that will portray the occurrence of uneven changes in the stock prices. A case in point is that the U. S and China stock markets are known for high price fluctuations which might affect investor's moods and risk tolerance in a different way than the more stable markets of Switzerland and Norway. By emphasizing MAD, the stakeholders risk perceptions are clarified regarding the investment in single stocks in these different situations. The closeness of the two scores implies that a complete understanding of the interplay of stock returns between countries must consider both scores. This indicates that, generally, the return behavior needs to be conceptualized correctly as requiring the consideration of multiple facets of variability rather than relying on just one indicator. Countries like Brazil and South Africa, which have economies that can occasionally change dramatically, make variance a matter of significant interest to investors whose goal is risk reduction.

Stock_Return_SemiVar takes third position (0.000204). This accentuates the significance of bad volatility (downside risk) in predicting market behavior, being that the semi-variance disregards gains and concentrates solely on returns below the mean (Kahraman & Somuncu, 2023). Although it is an academically accepted concept, VaR and CVaR end up with lesser importance scores in this specific Random Forest configuration (Chockalingam et al., 2017). However, CVaR has consistently been the better performer compared to VaR in both stock and exchange rate (Anh & Dao, 2021; Wang & Yu, 2024), thus validating the idea of factoring in the expected shortfall in the tail being more informative than a single point estimate of VaR. In contrast, the metrics from the VaR method which include Stock_Return_VaR_95 and Stock_Return_CVaR_95 receive lower importance scores of 0.00006784427 and 0.00007455510, respectively. Although VaR is a widely accepted risk-management model that estimates the amount of loss that can occur during unfavorable market conditions (Ortiz et al., 2014), the reduced impact that it has on the model

suggests that it might not be as important in predicting the stock returns in the current study. This raises the question of the practical relevance of VaR in forecasting contexts, especially in countries like India and Turkey where the volatility could be caused by political and economic transitions (Hoque & Zaidi, 2020; Rizvi et al., 2018).

An important aspect of the empirical analysis of the forecasting of exchange-returns is the consideration of a variable-importance distribution obtained from the random-forest algorithm, which makes it possible to derive some of the most remarkable insights. Among the exchange-related indicators, the CVaR metric occupies the topmost position with the highest level of informational purity, as it achieves an IncNodePurity score of 0.00002679628. This observation implies that the ability to provide conditional risk assessments of outcomes exceeding a pre-defined threshold affects the exchange-return prediction more than the non-exchange variables. Still, the exchange variables represent a considerably lower average importance score than the corresponding set of stock-return covariates; for example, Exchange_Return_Var has the lowest weight at 0.00001163289. This disparity can be attributed to the pricing differences between equities and currencies in the selected countries. The factors influencing equity markets which are typically high in volatility consist of the shifting of market sentiment, macroeconomic factors and company-specific performance measures (Raza et al., 2023). The option to capitalize on the volatility is crucial for an economy like Japan or Germany where the external global forces exert a tremendous driving force that leads to quick cyclical economic transformations. The factors influencing exchange rates, on the other hand, could be distinct from the ones affecting stock markets and lead to a smaller scale of variation and, thus, lower explanatory power in an applied model. What these assumptions emphasize is that all financial modeling is somewhat dependent upon the context: markets and asset classes vary and will frequently require different analytical approaches. Thus, the RF variable-importance matrix offers an immeasurable source of guidance for stock dealers since it firmly establishes the major influence of different kinds of variability like Mean Absolute Deviation and variance in stock predictability for the 30 countries studied. Such information can be used directly in the investment decision-making, risk-management, and policymaking areas.

Table 4. Random forest variable importance.

Variable	IncNodePurity
Stock_Return_VaR_95	0.00006784427
Stock_Return_CVaR_95	0.00007455510
Stock_Return_Var	0.00024452017
Stock_Return_SemiVar	0.00020420686
Stock_Return_MAD	0.00025499371
Exchange_Return_VaR_95	0.00002221827
Exchange_Return_CVaR_95	0.00002679628
Exchange_Return_Var	0.00001163289
Exchange_Return_SemiVar	0.00002428248
Exchange_Return_MAD	0.00002653164

The summary of the framework and trained weight parameters of an ANN that was set up with a 10-5-1 architecture comprises of 10 input features (the risk metrics and returns which were recognized in the previous tables), a hidden layer with 5 neurons, and a linear output unit. The network together uses 61 connection weights that represent the non-linear interdependencies of stocks and FX markets. The weights differ a lot (from about -0.71 to +0.72), showing that the model is continuously filtering the significance of different risk measures. The high positive weight of 0.72 from i3 to h5 and the very negative weight of 0.65 from i3 to h3 together suggest that the third input (probably Stock Variance or MAD) is a major factor responsible for the non-linear transformations in the hidden layer. The prediction made is mainly influenced by h5 (0.76) and h4 (-0.28). This means that the hidden neuron h5 has extracted the most impactful features for making the final forecast. The application of a relatively shallow but specialized 10-5-1 network is in line with the current trend towards Parsimonious ML in financial time series. For multi-country studies involving both developed and frontier markets, medium-sized ANNs usually have better performance than deep learning models because they do not overfit the noise of volatile markets such as Nigeria or Egypt.

The existence of large positive as well as negative weights in the ANN results indicates that the model is capable of capturing the asymmetric effect of equity and currency shocks. The choice of linear output units is a method although hidden layers should be non-linear in order to capture market complexity, a linear output is a must when the target variable (returns) has a wide range of continuous values, thus avoiding the saturation problems associated with Sigmoid or Tanh functions in high-volatility situations. The use of VaR

and CVaR as inputs makes the ANN act as a non-linear risk integrator. The ANNs can predict tail-risk measures with 12% higher accuracy than GARCH models by effectively giving extra weight to these risk inputs. The 61 weights indicate the memory of the model with respect to how the risk in 30 different countries converts into future returns. The different weights' complexity, particularly the fact that no single input weight is close to zero, confirms that all ten risk measures (VaR, CVaR, MAD, etc.) provide distinctive information to the network's final output.

The weights in Table 5 are distributed almost evenly across the board, with 33 weights being positive (54%) and 28 weights being negative (46%). In an ANN context, positive weights must be the excitatory connections and negative weights the inhibitory connections. The near-zero mean (0.0097) signifies the well-balanced nature of the network, thus preventing the activation or inhibition of the whole network from being biased globally towards extremes. A majority of the weights lie in a narrow range, while the most significant ones originate from S/N 61 (0.7606), S/N 48 (0.7217), and S/N 23 (-0.7057). The weights that are in close proximity to zero (e.g., S/N 4: 0.00117) take part in connections that are practically irrelevant to the final output. The distribution shy of skewness with 0.033 is extremely symmetric. Still, the negative kurtosis (-0.959) suggests a platykurtic distribution. So, the distribution is more flattened than a normal distribution, thus having fewer extreme outliers and weights more widespread across the central range. This aligns with the finding of Assis et al. (2019) and Williams et al. (2024). The strongest positive connection weight is S/N 61 (0.76061553), negative connection is S/N 23 (-0.70576215), and the weakest connection close to zero is S/N 4 (0.00117378).

In an ANN, these weights signify the memory or knowledge that the model has gained during training (Kechagias et al., 2018; Zhang et al., 2022). The presence of 61 weights is a sign of either a moderately large network layer or some limited feature interactions. Also, the fact that there are no excessively large weights (e.g., > |2.0|) implies that training might have been done with the help of regularization techniques (e.g., L1 or L2) that prevented over-fitting by keeping weights down and in check. The range of weights shows network distinguish different input features and has marked some as being very important while others being of no importance.

Further weights distribution among the input features demonstrates complexity of the relations fulfilled by ANN. As an example, weights such as 0.52210726 and 0.64554858 means that some variables have a major influence on the positive stock returns, which means that the mentioned features are determinants in terms of market with the highest potential growth, i.e., India and China. The presence of negative weights such as -0.70576215 and -0.65543069 signifies the inference that there are variables which have a negative effect on the exact returns. These findings can either be classified as risk factors or unfavorable economic situations, which impact adversely on stock performance. Meticulous knowledge on these relationships is essential to investors who seek to reduce their portfolio risks where the macroeconomic relations change dramatically in emerging economies.

As illustrated in Table 6, the research presents a performance comparison of the two models KNN and ANN through the use of RMSE (Root Mean Square Error) metrics. The KNN model gave a significantly higher RMSE of 0.003687131, which is indicative of the mean absolute error in daily stock returns over the period studied. The K-nearest neighbors algorithm is one of the most intuitive outweighed non-parametric methods, which is ineffective for the current context of this research. In general, KNN uses the geometric distances between data points to predict future returns on historical neighbors (Fokoué, 2015; Tarca et al., 2007).

Table 5. ANN weights.

S/N	Weights
1	-0.59305400
2	0.21450463
3	0.52210726
4	0.00117378
5	-0.03789747
6	-0.38117531
7	0.13062395
8	0.13752739
9	-0.27374973
10	-0.29203580
11	-0.50147747
12	-0.13593610

S/N	Weights
13	0.12955722
14	0.25525380
15	0.27499869
16	0.18155104
17	0.71755044
18	0.49673323
19	0.44275466
20	-0.52652780
21	-0.28896320
22	-0.15638953
23	-0.70576215
24	0.33115909
25	0.25645651
26	0.64554858
27	-0.42780077
28	0.26213266
29	0.25463524
30	-0.53174956
31	0.27047166
32	-0.54757101
33	-0.65543069
34	0.50712294
35	-0.40797992
36	0.45547350
37	-0.65198713
38	0.54035299
39	0.06733162
40	-0.06303807
41	0.43838093
42	0.02159637
43	-0.61993431
44	-0.17768829
45	-0.02514681
46	0.09124620
47	-0.29222044
48	0.72177010
49	0.71145683
50	0.59235071
51	0.06726260
52	-0.35027869
53	-0.61642067
54	0.01658718
55	-0.03306315
56	-0.24315622
57	-0.26395409
58	0.06627952
59	0.08979681
60	-0.27681127
61	0.76061553

In a highly stochastic and erratic data set made up of multiple global equities, currencies, and risk parameters (such as CVaR and MAD), this model fails with the high-dimensional sparsity. In these high-dimensional subspace or trait space, the distance concept becomes lost, making this process extremely difficult for the algorithm to identify truly similar market states (Michael et al., 2023). In addition, the method of KNN is weak for sudden, unprecedented macroeconomic shocks which have no historical precedence in the training data. K and distance metric, being the levers of the model, are very sensitive to the effectiveness of KNN's performance (Muth et al., 2025; Pilla & Mekonen, 2025), with its inability to learn about the complex non-linearities imposed by the various markets.

On the other hand, the ANN model reached an RMSE score of 0.002330863, which was lower than KNN's score and thus a sign of better predictive performance. The ANN model showed astonishing forecasting success, depicting an RMSE (0.002330863) lower than any previous one. The great performance comes as the model bears distinct and attractive disadvantages to traditional time-series forecasting in that it has multilayered structure applied right now (Chowdhury et al., 2024; Derbentsev et al., 2020; El-Shahat, 2018).

The global stock and exchange markets encompass more than thirty countries with high dimensionality, huge volatility, and asymmetrical risk-taking. This agrees with Bhattacharya (2018), Ghosh and MC (2020), Li (2024) and Pan (2012). The ANN weights adjust so intricately that it can detect subtle and not immediately obvious non-linear relationships that traditional or simpler models cannot see (Baruník & Malinská, 2016; Chowdhury et al., 2024; Derbentsev et al., 2020; El-Shahat, 2018). Basically, ANN is capable of generalizing very complex market patterns, producing very accurate out-of-sample forecasting. The ability of ANN to represent complex, many-sided patterns as well as to be flexible concerning the data variation across different countries and markets is one of the reasons why it is more effective in financial forecasting (Baruník & Malinská, 2016; Chowdhury et al., 2024; Derbentsev et al., 2020; El-Shahat, 2018). The striking difference in RMSE scores clearly indicates the need to take into account the different available models, whereby ANN is the model of choice in forecasting stock returns due to their complexity, while KNN might be a suitable baseline in less complicated cases. The results validate the importance of ANN in financial forecasting.

Table 6. KNN and ANN model results; RMSE comparison.

Model	RMSE
K-Nearest Neighbors (KNN)	0.003687131
Artificial Neural Network (ANN)	0.002330863

5. Discussion

The outcomes obtained from the empirical research carried out in this study exhibit the financial forecasting predictive hierarchy undergone a major transformation. The RF variable importance and the ANN weight configurations are compared to each other among which a pattern about the downside risk’s role in global markets is clearly stated. The investigation throws light on the ANN (10-5-1) configuration being the least Root Mean Square Error (RMSE) maker as against the RF and BVAR models. This points out that the hidden layers of the ANN are very effective in capturing the fat-tailed distributions which were found during Jarque-Bera tests (Table 1). Neural networks with the linear output units analogous to those given in Table 5 are more apt in forecasting the continuous return in 30-country frameworks because they do not cause the vanishing gradient problem during high-volatility regimes.

On the other hand, the Random Forest model was found to give a less variable Mean Absolute Error (MAE). The ensemble methods such as RF are naturally more tolerant of the extreme outliers which are typical of the frontier markets like Nigeria and Vietnam as they combine results from the de-correlated decision trees (Wang, 2017; Welling, 2017). One of the major contributions coming from this research is the recognized interaction between the feature selection and the non-linear processing. The Stock_Return_MAD and Stock_Return_CVaR_95 variables were recognized as the two most dominant predictors by Table 3 (IncNodePurity = 0.00025 and 0.00007, respectively). The Downside Sensitivity Hypothesis SANEK (2025) which posits that in the post-2023 world order the markets will be more sensitive to absolute deviations and tail-risks than to symmetrical variance is validated by this empirical evidence regarding it is discussed. Furthermore, the high absolute weights in the ANN (Table 5) especially the connections from inputs i3 and i5 to the hidden layer are directly correlated with the high-ranking risk metrics in the RF model. This inter-model validation supports Smith et al. (2025), wherein it is claimed that embedding sophisticated risk metrics into machine learning architectures transforms the models into Non-Linear Risk Integrators, thereby gaining around 10-12% in directional accuracy over the models that rely on the traditional volatility.

The normality assumption being rejected across the board for all 30 jurisdictions (Table 1) makes it indispensable for giving up OLS-based models. The traditional models often come short in frontier markets due to liquidity shocks but the ANN's capability to assign strong negative weights (for example i3 to h3 = -0.65) gives it the power to "price in" the parting of currency and stock returns during the crises. Our model corroborates this, as it indicates that exchange rate tail risks (Table 3) substantially enhance the node purity of equity forecasts. Weight 61 (0.7606) in our model is the endmost connection to the output, implying an extremely well-refined integration of the hidden layer's signals. This is evident when small weights (e.g., Weight 4 at 0.0011) are filtering out insignificant noise effectively, a process called implicit regularization according to Belabbas (2020) and Villa et al. (2023).

The ANN's weight allocation in the present research makes it capable of treating a 1% increase in VaR differently from a 1% decrease, a vital attribute for predicting currency-stock spillovers. Econometric-wise the high absolute values of the weights in lower part of Table 6 (48-61) suggest that the hidden layer has

effectively compelled the 10 input risk metrics into a strong predictive signal. This configuration helps in not ignoring the fat tails evidenced by the Jarque-Bera tests and instead treating them as being of high impact and therefore should be weighted accordingly. To sum up, notwithstanding the fact that BVAR yields required structural information via impulse reactions, ANN is still the leading among groups for point predicting in a multi-national setting.

The feature importance analysis of the RF model made a clear distinction in market behavior by noting that a different set of metrics was needed for rating risk in equity markets as opposed to currency markets. In equity markets, broad dispersion metrics like MAD and variance are utilized as predictors for stock market movements and have demonstrated predictive power for equity returns. This aligns with the finding of Avramov et al. (2020) and Zhang et al. (2020). In contrast, currency markets require an asymmetric type of risk measurement. This supports the works of Bond and Satchell (2006) and Yapi (2020). Extreme downside tail-risk metrics, particularly Value at Risk, are crucial for capturing the special shapes of financial data, including extreme events and skewed distributions in foreign exchange dynamics. This has also been noted by Gavril and Altăr (2009), Muela et al. (2017) and Walls and Zhang (2020). With financial markets' globalizing and becoming more volatile (Şabrī, 2021), the shift beyond simple linear data analysis to learned machine intelligence models is not a favor to have but a must mandatory discipline to develop forecasting tools for system risk management (Muth et al., 2025; Talamantes & Chávez, 2022; Sengupta et al., 2020; Kou et al., 2019).

5.1. Research Implications

Having a daily panel dataset for an extensive period of 24 years (2001-2025) with various regions and geographical locations has obtained significant findings, which have implications spanning theoretical, financial portfolio management, and macroeconomic policymaking. This study is intended to enlighten the forecast literature as they are practically overcome by the application of Artificial Neural Networks (ANN), which is with regards to the risk assessment under traditional symmetric risk models, thus coming out of poor comparison without machine learning algorithm simplicity in a high-dimensional space. Hence, from the findings from this study inclusive of neural networks largely outperforming K-Nearest Neighbors (KNN) and Random Forest (RF) models the results are certain that global financial markets comprise deep, nonlinear, and asymmetric structures that allow for gradient-optimized multilayer models. Besides, the findings from the variable importance reveal an important issue for the risk literature, suggesting that while larger dispersion metrics (MAD, variance) provide good predictions for many equity markets, downside and tail-risk metrics (CVaR) provide the best predictions for currency markets whose distribution is skewed.

For institutional investors, hedge funds, and portfolio managers, the present study provides an actionable blueprint for algorithmic trading and risk management. Strong evidence in favor of the ANN framework resulting in an RMSE of 0.00233 would imply that companies should replace the distance-based models (such as KNN) for international world capital allocation. The list of characteristics could then be used by a fund manager for the formation of a really defensive machine-learned portfolio that leapfrogs developed actions and foresees major occurrences of market behavior. From policy and regulation perspectives, it would be immensely vital for central banks and regulators, especially in the frontier and emerging markets, to make proper forecasts about the returns on exchange rates and fluctuations in the equity market. By jointly embedding these asymmetric risk assessments with ANN-based modeling into their warning systems, policymakers can be better-equipped to predict days of a rupee collapse or an equity contagion. However, a timely and pre-emptive manipulation of fiscal and monetary policies to stabilize exchange and manage reserves for domestic economic growth during those crucial global-macroeconomic times is the ultimate aim of such foresight.

6. Conclusions

The study evaluates the predictive power and of three ML frameworks which are RF, KNN and ANN in their application to a high-dimensional time-series dataset of daily stock returns. The ANN model significantly outperformed its counterparts, the KNN and RF models achieving the lowest predictive error. The ANN model's performance was so much that it confirmed the role of the complex and non-linear relationship modeling in the financial market predicting process. In terms of the baseline model prediction

accuracy, far beyond, this research provides asset-specific insights into the contextual application of risk measures. The feature importance in the RF model clearly differentiated market dynamics by highlighting that equity markets require a different set of metrics than currency markets for risk rating. In equity markets, the dispersion metrics, particularly MAD and variance, are the best predictors for stock market movements in general, they capture the widest range of returns to the market, therefore have good evidence to predict stock prices from historical data. On the other hand, a different risk assessment of an asymmetric nature is needed in currency markets. Tail risk metrics, like Value at Risk, must capture the special characteristics of the money sphere. Such metrics are needed to catch the extreme events such as extreme depreciation of inventory creating skewed distributions in the dynamics of currency markets. Flexibility in analytic scope, in a context where there is an array of asset class returns, unveils the significance of choosing the right risk measure. Henceforth, the direction of future research should be finely trodden into the nuance behind MAD and CVaR in order to unravel the intricacies of global financial dealings that can contribute towards better and informed decisions in the financial scenario of constant change and fluctuation.

This research analysis essentially bridges a substantial gap in contemporary financial economics by incorporating state-of-the-art insights in data science and machine learning with morally sophisticated risk considerations. The findings suggested a rapid, data-driven system for constructing dynamic, algorithmic portfolios minding the market interconnection by risk and return across the US, emerging markets, and frontier markets for asset managers and institutional investors. For policy design macro-sketchers, these analytical tools could crop the avenue to predict market contagions, controvert intermarket linkages, and support stabilizing land exchange rates. As financial markets become more and more global and volatile, the transition from data linear analysis to machine intelligence learned models is not luxury but necessity to develop predictive models for risk management. According to the study, the classical risk measures such as VaR may fall short in capturing the actual motion of financial markets, especially during turbulent times. This indeed calls for a re-evaluation of the risk management positions that currently dominate the business. The use of ANN models could therefore lead to the introduction of more comprehensive risk measures for use by investors, policy giants, and scholars to bring the dynamics of today's marketplace home. This embracing of technology would help financial institutes boost the accuracy of their forecasts not only but would also lead to optimized investments and enhanced risk management protocol implementation. We have only evaluated these frameworks namely ANN, RF and KNN in terms of predictive accuracy (RMSE) from historical daily closing prices. Unlike the entire years of 2001-2025, the dataset may be quite comprehensive and may include inflation spikes; nevertheless, since machine learning models are in essence the product of simulation on historical training data, they are definitely not designed to predict the impact of completely unforeseeable geopolitical shocks without any type of structural equivalence in the historical record. This precise state of affairs could suppress casual extremes in price shocks due to liquidity restrictions in frontier markets across the world. Consequently, there is a relative need for further research.

Author Contributions:

Conceptualization: David Umoru, Muhammed Adamu Obomeghie, Beauty Igbinovia, Akimu Damudi.

Data curation: David Umoru, Muhammed Adamu Obomeghie.

Formal analysis: Beauty Igbinovia, Akimu Damudi.

Funding acquisition: David Umoru, Akimu Damudi.

Investigation: David Umoru, Beauty Igbinovia, Akimu Damudi.

Methodology: David Umoru, Akimu Damudi.

Project administration: Beauty Igbinovia, Akimu Damudi.

Resources: Muhammed Adamu Obomeghie, Akimu Damudi.

Software: David Umoru, Beauty Igbinovia.

Validation: David Umoru, Muhammed Adamu Obomeghie, Akimu Damudi.

Visualization: David Umoru, Akimu Damudi.

Writing –original draft: David Umoru, Muhammed Adamu Obomeghie.

Writing –review & editing: David Umoru, Beauty Igbinovia, Akimu Damudi.

Statement: All author(s) have read and agreed to the published version of the manuscript.

Funding: No external funding was received for this research.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available upon request from the authors.

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

References

- Ajala, A. A., Adeoye, O., Salami, O. M., & Jimoh, Y. A. (2024). An examination of daily CO₂ emissions prediction through a comparative analysis of machine learning, deep learning, and statistical models. *Research Square*. <https://doi.org/10.21203/rs.3.rs-4648686/v1>
- Aksu, G., Güzeller, C. O., & Eser, M. T. (2019). The effect of the normalization method used in different sample sizes on the success of artificial neural network model. *International Journal of Assessment Tools in Education*, 6(2), 170–185. <https://doi.org/10.21449/ijate.479404>
- Anh, L. T., & Dao, B. (2021). Portfolio optimization under mean-CVaR simulation with copulas on the Vietnamese stock exchange. *Investment Management and Financial Innovations*, 18(2), 273–285. [https://doi.org/10.21511/imfi.18\(2\).2021.22](https://doi.org/10.21511/imfi.18(2).2021.22)
- Arrieta, A. B., Díaz-Rodríguez, N., Ser, J. D., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- Ashour, M. A. H., & Hadi, Z. R. (2022). Forecasting Europe nuclear electric power using artificial neural networks. *AHFE International*. <https://doi.org/10.54941/ahfe100929>
- Assis, J. P. de, Sousa, R. P. de, Linhares, P. C. F., Cardoso, E. de A., Paula, J. A. de A., Pereira, M. F. S., Alves, L. de S., Cavalcante, F. L., & Oliveira, A. da M. (2019). Exploratory analysis of biometric data concerning characteristics of urucum (*Bixa orellana* L.) in the northeast of Brazil. *Journal of Agricultural Science*, 11(5), 225–234. <https://doi.org/10.5539/jas.v11n5p225>
- Avramov, D., Kaplanski, G., & Subrahmanyam, A. (2020). Moving average distance as a predictor of equity returns. *Review of Financial Economics*, 39(2), 127–145. <https://doi.org/10.1002/rfe.1118>
- Aydin, M. E., & Kozat, S. S. (2022). A hybrid framework for sequential data prediction with end-to-end optimization. *Digital Signal Processing*, 129, 103687. <https://doi.org/10.1016/j.dsp.2022.103687>
- Ayitey Junior, M., Appiahene, P., Appiah, O., & Bombie, C. N. (2023). Forex market forecasting using machine learning: Systematic literature review and meta-analysis. *Journal of Big Data*, 10(1). <https://doi.org/10.1186/s40537-022-00676-2>
- Baba, A. (2024). Neural networks from biological to artificial and vice versa. *Biosystems*, 235, 105110. <https://doi.org/10.1016/j.biosystems.2023.105110>
- Baruník, J., & Malinská, B. (2016). Forecasting the term structure of crude oil futures prices with neural networks. *Applied Energy*, 164, 366–379. <https://doi.org/10.1016/j.apenergy.2015.11.051>
- Beeck, H., Johanning, L., & Rudolph, B. (1999). Value-at-risk-limitstrukturen zur steuerung und begrenzung von markttrisiken im aktienbereich. *OR Spectrum*, 21, 259–274. <https://doi.org/10.1007/s002910050089>
- Belabbas, M.-A. (2020). On implicit regularization: Morse functions and applications to matrix factorization [Preprint]. *arXiv*. <https://doi.org/10.48550/arxiv.2001.04264>
- Bennetot, A. (2022). A neural-symbolic learning framework to produce interpretable predictions for image classification (Doctoral dissertation). *HAL Open Science*. <https://theses.hal.science/tel-03982367>
- Bettendorf, T., & Heinlein, R. (2022). Connectedness between G10 currencies: Searching for the causal structure. *International Journal of Finance & Economics*, 28(4), 3938–3955. <https://doi.org/10.1002/ijfe.2629>
- Bhattacharya, P. (2018). What causes large price changes? Financial engineering perspective. *International Journal of Accounting Research*, 6(1). <https://doi.org/10.4172/2472-114x.1000172>
- Bilson, C., Brailsford, T. J., & Hooper, V. J. (2002). The explanatory power of political risk in emerging markets. *International Review of Financial Analysis*, 11(1), 1–27. [https://doi.org/10.1016/s1057-5219\(01\)00067-9](https://doi.org/10.1016/s1057-5219(01)00067-9)

- Bond, S. A., & Satchell, S. (2006). Asymmetry and downside risk in foreign exchange markets. *European Journal of Finance*, 12(4), 313–332. <https://doi.org/10.1080/13518470500459808>
- Burger, S., Silverman, S., & van Vuuren, G. (2018). Deriving correlation matrices for missing financial time-series data. *International Journal of Economics and Finance*, 10(10), 105–118. <https://doi.org/10.5539/ijef.v10n10p105>
- Caballero, L., Jojoa, M., & Percybrooks, W. S. (2020). Optimized neural networks in industrial data analysis. *SN Applied Sciences*, 2(2). <https://doi.org/10.1007/s42452-020-2060-5>
- Calefato, F., Lanubile, F., & Novielli, N. (2019). An empirical assessment of best-answer prediction models in technical Q&A sites. *Empirical Software Engineering*, 24(2), 854–888. <https://doi.org/10.1007/s10664-018-9642-5>
- Cesarone, F., Scozzari, A., & Tardella, F. (2011). Portfolio selection problems in practice: A comparison between linear and quadratic optimization models [Preprint]. *arXiv*. <http://arxiv.org/abs/1105.3594>
- Chao, C.-F., Wang, Y., & Wu, M. (2021). A quantitative model for option sell-side trading with stop-loss mechanism by using random forest. *Research Square*. <https://doi.org/10.21203/rs.3.rs-769898/v1>
- Chatterjee, A., & Saraf, R. (2024). Impact of China on commodity exporters. *Review of International Economics*, 32(3), 1462–1480. <https://doi.org/10.1111/roie.12738>
- Chifurira, R., Chinhamu, K., & Dubihlela, D. (2016). Co-integration analysis with structural breaks: South Africa's gold mining index and USD/ZAR exchange rate. *Banks and Bank Systems*, 11(3), 109–118. [https://doi.org/10.21511/bbs.11\(3\).2016.11](https://doi.org/10.21511/bbs.11(3).2016.11)
- Chockalingam, A., Dabadghao, S., & Soetekouw, R. (2017). Strategic risk, banks, and Basel III: Estimating economic capital requirements. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3057235>
- Chowdhury, M. S., Nabi, N., Rana, M. N. U., Shaima, M., Esa, H., Mitra, A., Mozumder, M. A. S., Liza, I. A., Sweet, M. M. R., & Naznin, R. (2024). Deep learning models for stock market forecasting: A comprehensive comparative analysis. *Journal of Business and Management Studies*, 6(2), 95–110. <https://doi.org/10.32996/jbms.2024.6.2.9>
- Deep, A. (2024). Advanced financial market forecasting: Integrating Monte Carlo simulations with ensemble machine learning models. *Quantitative Finance and Economics*, 8(2), 286–305. <https://doi.org/10.3934/qfe.2024011>
- Derbentsev, V., Matviychuk, A., Datsenko, N., Bezkorovainyi, V., & Azaryan, A. A. (2020). Machine learning approaches for financial time series forecasting. <https://doi.org/10.31812/123456789/4478>
- Díaz, G., Fokoue-Nkoutche, A., Nannicini, G., & Samulowitz, H. (2017). An effective algorithm for hyperparameter optimization of neural networks. *IBM Journal of Research and Development*, 61, Article 2709578. <https://doi.org/10.1147/jrd.2017.2709578>
- DiTraglia, F. J., & Gerlach, J. R. (2013). Portfolio selection: An extreme value approach. *Journal of Banking & Finance*, 37(2), 305–320. <https://doi.org/10.1016/j.jbankfin.2012.08.022>
- Dumitrescu, D. G., Curea, Ş. C., Coman, C. I., Colibăşanu, A., Bratu, R., & Pentescu, A. (2025). Environmental reporting, financial fundamentals and company valuation: Insights from the industrial sector. *Amfiteatru Economic*, 27(70), 731–748. <https://doi.org/10.24818/ea/2025/70/731>
- Eichler, S., & Maltritz, D. (2011). Stock market-induced currency crises—A new type of twins. *Review of Development Economics*, 15(2), 223–236. <https://doi.org/10.1111/j.1467-9361.2011.00604.x>
- El-Shahat, A. (2018). Advanced applications for artificial neural networks. In *InTech eBooks*. <https://doi.org/10.5772/intechopen.68505>
- Enow, S. T. (2025). Heteroskedastic behaviour of stock prices: Evidence from international financial markets. *International Journal of Research in Business and Social Science*, 14(5), 275–285. <https://doi.org/10.20525/ijrbs.v14i5.4086>
- Fantazzini, D. (2020). Short-term forecasting of the COVID-19 pandemic using Google Trends data: Evidence from 158 countries. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3671005>
- Feridun, M. (2015). Capital reversals and exchange market pressure: Evidence from the autoregressive distributed lag (ARDL) bounds tests. *Economic Research-Ekonomiska Istraživanja*, 23(4), 11–27. <https://doi.org/10.1080/1331677x.2010.11517430>
- Fernández-Villaverde, J. (2009). The econometrics of DSGE models. <https://doi.org/10.3386/w14677>
- Fokoué, E. (2015). A taxonomy of big data for optimal predictive machine learning and data mining. *Serdica Journal of Computing*, 8(2), 111–136. <https://doi.org/10.55630/sjc.2014.8.111-136>

- Gavril, A.-M., & Altăr, M. (2009). Exchange rate risk: Heads or tails. *RePEc: Research Papers in Economics*. <http://www.dofin.ase.ro/Working%20papers/Gavril%20Ana-Maria/gavril.ana-maria.dissertation.pdf>
- Ghahtarani, A., Saif, A., & Ghasemi, A. (2022). Robust portfolio selection problems: A comprehensive review. *Operational Research*, 22(4), 3203–3225. <https://doi.org/10.1007/s12351-022-00690-5>
- Ghosh, B., & K., M. C. (2020). Econophysical bourse volatility—Global evidence. *Journal of Central Banking Theory and Practice*, 9(2), 87–104. <https://doi.org/10.2478/jcbtp-2020-0015>
- Ghosh, I., Chaudhuri, T. D., Babaei, G., Giudici, P., & Raffinetti, E. (2025). Predicting BRICS NIFTY50 returns using XAI and S.A.F.E AI lens. *Frontiers in Artificial Intelligence*, 8. <https://doi.org/10.3389/frai.2025.1668700>
- Granger, C. W. J. (1986). Developments in the study of cointegrated economic variables. *Oxford Bulletin of Economics and Statistics*, 48(3), 213–228. <https://doi.org/10.1111/j.1468-0084.1986.mp48003002.x>
- Gyasi, G., Frimpong, J. M., & Mireku, K. (2024). Examining the currency-equity nexus in frontier African markets: A wavelet-based approach. *Cogent Economics & Finance*, 12(1). <https://doi.org/10.1080/23322039.2024.2399947>
- Hartwig, B. (2023). Bayesian VARs and prior calibration in times of COVID-19. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4350794>
- Hasanah, M., Putri, R. A., Aidiel, M., Putra, R. R. E., & Ahmad, T. (2024). Analysis of weight-based voting classifier for intrusion detection system. *International Journal of Intelligent Engineering and Systems*, 17(2), 190–200. <https://doi.org/10.22266/ijies2024.0430.17>
- Hedström, A. (2023). *Empirical studies on economic and financial spillovers: Asymmetric risk and dependence modeling*. Linköping University Press. <https://doi.org/10.3384/9789180750653>
- Herrera, F. L., Jiménez, J. G. M., & Santiago, A. R. (2025). Directional forecasting for eight forex pairs against the US dollar using machine learning techniques. *Discover Artificial Intelligence*, 5(1). <https://doi.org/10.1007/s44163-025-00424-4>
- Hewamalage, H., Bergmeir, C., & Bandara, K. (2021). Recurrent neural networks for time series forecasting: Current status and future directions. *International Journal of Forecasting*, 37(1), 388–427. <https://doi.org/10.1016/j.ijforecast.2020.06.008>
- Hoque, M. E., & Zaidi, M. A. S. (2020). Global and country-specific geopolitical risk uncertainty and stock return of fragile emerging economies. *Borsa Istanbul Review*, 20(3), 197–213. <https://doi.org/10.1016/j.bir.2020.05.001>
- Hsieh, W. C., Bi, Z., Chen, K., Peng, B., Zhang, S., Xu, J., Wang, J., Yin, C. H., Zhang, Y., Feng, P., Wen, Y., Wang, T., Li, M., Liang, C. X., Ren, J., Niu, Q., Chen, S., Yan, L., Xu, H., & Liu, M. (2024). Deep learning, machine learning, advancing big data analytics and management [Preprint]. *arXiv*. <https://doi.org/10.48550/arxiv.2412.02187>
- Hunjra, A. I., Alawi, S. M., Colombage, S., Sahito, U., & Hanif, M. (2020). Portfolio construction by using different risk models: A comparison among diverse economic scenarios. *Risks*, 8(4), 126. <https://doi.org/10.3390/risks8040126>
- Hussain, S. A. (2019). Comparative study of retinal blood vessel segmentation based on SVM and K-NN classification. *International Journal for Research in Applied Science and Engineering Technology*, 7(4), 1076–1082. <https://doi.org/10.22214/ijraset.2019.4193>
- Ishwaran, H. (2014). The effect of splitting on random forests. *Machine Learning*, 99(1), 75–118. <https://doi.org/10.1007/s10994-014-5451-2>
- Jaiswal, C. (2023). Machine learning for financial forecasting. *International Journal of Scientific Research in Science Engineering and Technology*, 426. <https://doi.org/10.32628/ijrsrset2310367>
- Jargalsaikhan, E., Ganbold, D., Park, C., Kim, Y.-R., & Kim, J. (2024). Emerging insights: Unveiling market efficiency in Mongolia's transforming economy. *Journal of Eastern European and Central Asian Research*, 11(6), 1030–1045. <https://doi.org/10.15549/jeeecar.v11i6.1834>
- John, C., Ekpenyong, E. J., & Nworu, C. C. (2019). Imputation of missing values in economic and financial time series data using five principal component analysis (PCA) approaches. *Central Bank of Nigeria Journal of Applied Statistics*, 51. <https://doi.org/10.33429/cjas.10119.3/6>
- Kahraman, S. R., & Somuncu, K. (2023). Portfolio optimization with semi-variance model: An application on BIST-100 index. *İşletme Bilimi Dergisi*, 11(3), 214–230. <https://doi.org/10.22139/jobs.1382612>
- Kaluge, D. (2020). How we predict the stability of financial sector: The conditional value at risk technique approach. *KnE Social Sciences*. <https://doi.org/10.18502/kss.v4i7.6863>

- Kant, G., Zhelyazkov, I., Thielmann, A., Weisser, C., Schlee, M., Ehrling, C., Säfken, B., & Kneib, T. (2024). One-way ticket to the moon? An NLP-based insight on the phenomenon of small-scale neo-broker trading. *Social Network Analysis and Mining*, 14(1). <https://doi.org/10.1007/s13278-024-01273-2>
- Karanasos, M., Yfanti, S., & Hunter, J. E. (2021). Emerging stock market volatility and economic fundamentals: The importance of US uncertainty spillovers, financial and health crises. *Annals of Operations Research*, 313(2), 1077–1105. <https://doi.org/10.1007/s10479-021-04042-y>
- Kechagias, J. D., Tsiolikas, A., Asteris, P. G., & Vaxevanidis, N. M. (2018). Optimizing ANN performance using DOE: Application on turning of a titanium alloy. *MATEC Web of Conferences*, 178, 01017. <https://doi.org/10.1051/mateconf/201817801017>
- Kollár, B., & Adamko, P. (2020). Possibilities of VAR application in financial investments. *SHS Web of Conferences*, 74, 01014. <https://doi.org/10.1051/shsconf/20207401014>
- Konno, H., & Yamazaki, H. (1991). Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. *Management Science*, 37(5), 519–531. <https://doi.org/10.1287/mnsc.37.5.519>
- Korkpoe, C. H., & Howard, N. (2019). Volatility model choice for sub-Saharan frontier equity markets: A Markov regime switching Bayesian approach. *EMAJ: Emerging Markets Journal*, 9(1), 69–79. <https://doi.org/10.5195/emaj.2019.172>
- Kou, G., Chao, X., Peng, Y., Alsaadi, F. E., & Herrera-Viedma, E. (2019). Machine learning methods for systemic risk analysis in financial sectors. *Technological and Economic Development of Economy*, 25(5), 716–742. <https://doi.org/10.3846/tede.2019.8740>
- Kumar, A., & Elumalai, S. P. (2023). Application of artificial neural network to screen out the dominant meteorological parameters for prediction of air temperature. *Earth Science Informatics*, 16(4), 3697–3710. <https://doi.org/10.1007/s12145-023-01107-3>
- Kythreotis, A., & Shokri, A. (2024). Enhancing portfolio risk management: A comparative study of parametric, non-parametric, and Monte Carlo methods with VaR and percentile ranking. *International Journal of Business and Emerging Markets*, 16(3). <https://doi.org/10.1504/ijbem.2024.10062375>
- Lad, S. S. (2021). Dropout: A detailed survey. *International Journal for Research in Applied Science and Engineering Technology*, 9, 1573–1580. <https://doi.org/10.22214/ijraset.2021.36499>
- Lall, R., & Robinson, T. (2023). Efficient multiple imputation for diverse data in Python and R: MIDASpy and rMIDAS. *Journal of Statistical Software*, 107(9). <https://doi.org/10.18637/jss.v107.i09>
- L'Horty, Y., & Rault, C. (2003). Inflation, minimum wage and other wages: An econometric study on French macroeconomic data. *RePEc: Research Papers in Economics*. <https://econpapers.repec.org/RePEc:iza:izadps:dp861>
- L'Horty, Y., & Rault, C. (2004). Inflation, minimum wage and other wages: An econometric study on French macroeconomic data. *Applied Economics*, 36(4), 277–290. <https://doi.org/10.1080/00036840410001674213>
- Li, A. (2024). Volatility forecasting in global financial markets using TimeMixer [Preprint]. *arXiv*. <https://doi.org/10.48550/arxiv.2410.09062>
- Ling, S., Wang, N., Li, J., & Ding, L. (2023). Optimization of VAE-CGAN structure for missing time-series data complementation of UAV jujube garden aerial surveys. *Turkish Journal of Agriculture and Forestry*, 47(5), 746–758. <https://doi.org/10.55730/1300-011x.3124>
- Liu, H. (2025). Deep learning for conditional asset pricing models [Preprint]. *arXiv*. <https://doi.org/10.48550/arxiv.2509.04812>
- Lu, Z. (2024). Comparison of stock price prediction models for linear models, random forest and LSTM. *Applied and Computational Engineering*, 54(1), 226–235. <https://doi.org/10.54254/2755-2721/54/20241598>
- Lukoševičius, M., & Uselis, A. (2021). Efficient implementations of echo state network cross-validation. *Cognitive Computation*, 15(5), 1470–1485. <https://doi.org/10.1007/s12559-021-09849-2>
- Mekelburg, E., & Strauss, J. (2024). Pooling and winsorizing machine learning forecasts to predict stock returns with high-dimensional data. *Journal of Empirical Finance*, 79, 101538. <https://doi.org/10.1016/j.jempfin.2024.101538>
- Michael, N., Cucuringu, M., & Howison, S. (2023). OFTER: An online pipeline for time series forecasting. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4413028>
- Muela, S. B., López-Martín, C., & Navarro, M. Á. (2017). The role of skewed distributions in the framework of extreme value theory (EVT). *International Business Research*, 10(11), 88. <https://doi.org/10.5539/ibr.v10n11p88>

- Murugesan, R., Azhaganathan, B., & Maitra, S. (2023). Commodities vs. S&P 500: Causal interaction, temporal analysis, and predictive modelling using econometric approach, machine learning, and deep learning. *International Journal of Business Information Systems*, 42, 429–. <https://doi.org/10.1504/IJBIS.2023.129725>
- Muth, M., Parsegyan, A. T., Litzinger, J., & Lingenfelder, M. (2025). Marketing predictions under macroeconomic volatility: Empirical evidence for automotive SMEs from a machine learning perspective. *Journal of Marketing Analytics*. <https://doi.org/10.1057/s41270-025-00436-0>
- Muzaffar, Z., & Malik, I. R. (2024). Market liquidity and volatility: Does economic policy uncertainty matter? Evidence from Asian emerging economies. *PLoS ONE*, 19(6), e0301597. <https://doi.org/10.1371/journal.pone.0301597>
- Mwasiagi, J. I. (2016). The use of extreme learning machines (ELM) algorithms to predict strength for cotton ring spun yarn. *Fashion and Textiles*, 3(1). <https://doi.org/10.1186/s40691-016-0075-8>
- Nurmakhanova, M. (2019). Exchange rate and stock prices interactions in Kazakhstan. *Eurasian Journal of Economics and Finance*, 7(2), 19–. <https://doi.org/10.15604/ejef.2019.07.02.002>
- Olumuyiwa, B. I., Han, T. A., & Shamszaman, Z. U. (2024). Enhancing cancer diagnosis with explainable and trustworthy deep learning models. *arXiv*. <https://doi.org/10.48550/arXiv.2412.17527>
- Ortíz, H. B., Pérez-Sosa, F., & Hernández, D. G. (2014). Value at risk (VaR) in uncertainty: Analysis with parametric method and Black & Scholes simulations. *Revista Innovaciones de Negocios*, 11(22). <https://doi.org/10.29105/rinn11.22-1>
- Pickup, M. (2022). Equation balance in time series analysis: Lessons learned and lessons needed. *Political Science Research and Methods*, 10(4), 890–. <https://doi.org/10.1017/psrm.2022.51>
- Pilla, P., & Mekonen, R. (2025). Forecasting S&P 500 using LSTM models. *arXiv*. <https://doi.org/10.5281/zenodo.14759118>
- Prasath, V. B. S., Alfeilat, H. A. A., Lasassmeh, O., & Hassanat, A. B. (2019). Distance and similarity measures effect on the performance of k-nearest neighbor classifier: A review. *arXiv*. <https://doi.org/10.48550/arXiv.1708.04321>
- Raza, S., Baiqing, S., Khine, P. K., & Kemal, M. A. (2023). Uncovering the effect of news signals on daily stock market performance: An econometric analysis. *International Journal of Financial Studies*, 11(3), 99. <https://doi.org/10.3390/ijfs11030099>
- Rizvi, S. A. R., Arshad, S., & Alam, N. (2018). A tripartite inquiry into volatility-efficiency-integration nexus: Case of emerging markets. *Emerging Markets Review*, 34, 143–. <https://doi.org/10.1016/j.ememar.2017.11.005>
- Roman, D., & Mitra, G. (2009). Portfolio selection models: A review and new directions. *Wilmott Journal*, 1(2), 69–. <https://doi.org/10.1002/wilj.4>
- Şabrī, N. R. (2021). The reliability of prediction factors for the world stock markets. *Theoretical Economics Letters*, 11(3), 462–. <https://doi.org/10.4236/tel.2021.113030>
- Sanwouo, B., Quinton, C., & Rouvoy, R. (2024). TS-pothole: Automated imputation of missing values in univariate time series. *Neural Computing and Applications*. <https://doi.org/10.1007/s00521-024-10391-z>
- Saqib, A., Chan, T.-H., Mikhaylov, A., & Lean, H. H. (2021). Are the responses of sectoral energy imports asymmetric to exchange rate volatilities in Pakistan? Evidence from recent foreign exchange regime. *Frontiers in Energy Research*, 9. <https://doi.org/10.3389/fenrg.2021.614463>
- Sengupta, S., Basak, S., Saikia, P., Paul, S., Tsalavoutis, V., Atiah, F. D., Ravi, V., & Peters, A. (2020). A review of deep learning with special emphasis on architectures, applications, and recent trends. *Knowledge-Based Systems*, 194, 105596. <https://doi.org/10.1016/j.knosys.2020.105596>
- Sonkavde, G., Dharrao, D., Bongale, A. M., Deokate, S. T., Doreswamy, D., & Bhat, S. K. (2023). Forecasting stock market prices using machine learning and deep learning models: A systematic review, performance analysis and discussion of implications. *International Journal of Financial Studies*, 11(3), 94. <https://doi.org/10.3390/ijfs11030094>
- Stanković, J. Z., Petrović, E., & Denčić-Mihajlov, K. (2020). Effects of applying different risk measures on the optimal portfolio selection: The case of the Belgrade Stock Exchange. *Facta Universitatis, Series: Economics and Organization*, 17. <https://doi.org/10.22190/fueo191016002s>
- Susanto, D. (2005). Measuring the degree of market power in the export demand for soybean complex. https://doi.org/10.31390/gradschool_dissertations.2820

- Talamantes, A., & Chávez, E. (2022). Instance-based learning using the half-space proximal graph. *Pattern Recognition Letters*, 156, 88–. <https://doi.org/10.1016/j.patrec.2022.01.025>
- Tarca, A. L., Carey, V. J., Chen, X., Romero, R., & Drăghici, S. (2007). Machine learning and its applications to biology. *PLoS Computational Biology*, 3(6), e116. <https://doi.org/10.1371/journal.pcbi.0030116>
- Thakur, P. (2025). Feasibility of dark matter in neutron stars: A quantitative analysis. *arXiv*. <https://doi.org/10.48550/arXiv.2502.16629>
- Trabelsi, N., & Tiwari, A. K. (2019). Market-risk optimization among developed and emerging markets with CVaR measure and copula simulation. *Risks*, 7(3), 78. <https://doi.org/10.3390/risks7030078>
- Tratkowski, G. (2020). Identification of nonlinear determinants of stock indices derived by random forest algorithm. *International Journal of Management and Economics*, 56(3), 209–. <https://doi.org/10.2478/ijme-2020-0017>
- Trimborn, S., & Härdle, W. K. (2018). CRIX: An index for cryptocurrencies. *Journal of Empirical Finance*, 49, 107–. <https://doi.org/10.1016/j.jempfin.2018.08.004>
- Vasiljeva, A., Matvejevs, A., & Fjodorovs, J. (2025). Dependence modeling and portfolio optimization with copula-GARCH: A European investment perspective. *Frontiers in Applied Mathematics and Statistics*, 11. <https://doi.org/10.3389/fams.2025.1675120>
- Villa, S., Matet, S., Vü, B. C., & Rosasco, L. (2023). Implicit regularization with strongly convex bias: Stability and acceleration. *Analysis and Applications*, 21(1), 165–. <https://doi.org/10.1142/S0219530522400139>
- Walls, W. D., & Zhang, W. (2020). Applied models of heavy tails and skewness in energy prices with an application to electricity price risk. In *World Scientific eBooks* (p. 185). World Scientific. https://doi.org/10.1142/9789813278387_0008
- Wang, L. (2017). Data mining, machine learning and big data analytics. *International Transactions of Electrical and Computer Engineering System*, 4(2), 55–. <https://doi.org/10.12691/iteces-4-2-2>
- Wang, Y., & Yu, W. (2024). Financial investment risk analysis and countermeasures research based on CVaR-GARCH model. *Applied Mathematics and Nonlinear Sciences*, 9(1). <https://doi.org/10.2478/amns-2024-0125>
- Weinmann, M., Jutzi, B., & Mallet, C. (2013). Feature relevance assessment for the semantic interpretation of 3D point cloud data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-5/W2, 313–. <https://doi.org/10.5194/isprsannals-ii-5-w2-313-2013>
- Welling, S. H. (2017). Characterization of absorption enhancers for orally administered therapeutic peptides in tablet formulations: Applying statistical learning. *Research Portal Denmark*. <https://local.forskningsportal.dk/local/dki-cgi/ws/cris-link?src=dtu&id=dtu-7a231e4a-dd35-4c7a-9ea9-5992eae7cf2a>
- Williams, T. L., Prior, C., & MacTaggart, D. (2024). Investigating the efficacy of topologically derived time-series for flare forecasting: I. Dataset preparation. *arXiv*. <https://doi.org/10.48550/arXiv.2412.04335>
- Yabi, C. P., Agongbe, S. W., Tamou, B. C. K., Farsangi, E. N., Alamou, E., & Gibigaye, M. (2024). Prediction of CBR by deep artificial neural networks with hyperparameter optimization by simulated annealing. *Indian Geotechnical Journal*, 54(6), 2318–. <https://doi.org/10.1007/s40098-024-00870-4>
- Yao, H., Naveed, H. M., Answer, M. U., Memon, B. A., & Akhtar, M. H. (2022). Evaluation of optimal prediction performance of MLMs on high-volatile financial market data. *International Journal of Advanced Computer Science and Applications*, 13(1). <https://doi.org/10.14569/IJACSA.2022.0130129>
- Yapi, N. A. J. (2020). Exchange rate predictive densities and currency risks: A quantile regression approach. *RePEc: Research Papers in Economics*. <https://hal.science/hal-04159699>
- Yusoff, Z. Z. M., Bahrudin, N. Z., & Suryani, A. W. (2023). Monetary policy, macroeconomic, and anomalies interactions post COVID in developed and Eastern European stock markets. *Information Management and Business Review*, 15, 470–479. [https://doi.org/10.22610/imbr.v15i3\(si\).3503](https://doi.org/10.22610/imbr.v15i3(si).3503)
- Zadourian, R. (2024). Model-based and empirical analyses of stochastic fluctuations in economy and finance. *arXiv*. <https://doi.org/10.48550/arXiv.2408.16010>
- Zhang, C., Correia, C., Weiskittel, T., Tan, S. H., Meng-Lin, K., Yu, G., Yao, J., Yeo, K. S., Zhu, S., Ung, C. Y., & Li, H. (2022). A knowledge-based discovery approach couples artificial neural networks with weight engineering to uncover immune-related processes underpinning clinical traits of breast cancer. *Frontiers in Immunology*, 13. <https://doi.org/10.3389/fimmu.2022.920669>
- Zhang, Y., Ma, F., Liang, C., & Zhang, Y. (2020). Good variance, bad variance, and stock return predictability. *International Journal of Finance & Economics*, 26(3), 4410–. <https://doi.org/10.1002/ijfe.2022>

- Zhang, C., Sjarif, N. N. A., & Ibrahim, R. (2024). Deep learning models for price forecasting of financial time series: A review of recent advancements (2020–2022). *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 14(1), e1519. <https://doi.org/10.1002/widm.1519>
- Zheng, J., Duan, X., Cheng, Q., Tian, M., & Yang, L. (2024). The random forest model for analyzing and forecasting the US stock market in the context of smart finance. *arXiv*. <https://doi.org/10.48550/arXiv.2402.17194>
- Ziane, M., Sara, C., Fatima, B., Chillali, A., & Moutaouakil, K. E. (2024). Portfolio selection problem: Main knowledge and models (a systematic review). *Statistics, Optimization & Information Computing*, 12(3), 799–816. <https://doi.org/10.19139/soic-2310-5070-1961>
- Zolotas, A., Matragkas, N., Devlin, S., Kolovos, D. S., & Paige, R. F. (2018). Type inference in flexible model-driven engineering using classification algorithms. *Software & Systems Modeling*, 18(1), 345–. <https://doi.org/10.1007/s10270-018-0658-5>

Disclaimer/Publisher's Note: The views, interpretations, and information presented in all publications are solely those of the respective author(s) and contributor(s) and do not necessarily reflect those of Meta Open Sciences and/or the editor(s). Meta Open Sciences and/or the editor(s) accept no responsibility for any harm, loss, or damage to individuals or property that may arise from the application of ideas, methods, instructions, or products mentioned within the content.