BDDK Bankacılık ve Finansal Piyasalar
Journal of BRSA Banking and Financial Markets

Cilt / Volume: 18
Sayı / Issue: 2
Yıl / Year: 2024

e-ISSN: 1307-945X
Araştırma Makalesi
Research Article

Predictive Abilities of Machine Learning and Deep Learning Approaches for Exchange Rate Prediction

Furkan TÜRKOĞLU *
Eda GÖÇECEK **
Yavuz YUMRUKUZ ***

Abstract

This study evaluates the efficacy of forecasting models in predicting USD/TRY exchange rate fluctuations. We assess Support Vector Machine (SVM), XGBoost, Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU) models with 96 and 21 feature sets. Data from 01.01.2010 to 30.04.2024 were sourced from Bloomberg, CBRT, and BDDK. Findings indicate that LSTM and GRU models outperform traditional models, with GRU showing the highest predictive accuracy. SVM performs poorly with high-dimensional data, while XGBoost offers moderate predictive power but lacks in capturing intricate patterns. This study highlights the importance of model and feature selection in financial time series forecasting and underscores the advantages of advanced neural networks. The results provide valuable insights for analysts and policymakers in developing robust economic forecasting models.

Keywords: Exchange Rate, Machine Learning, Deep Learning, Time Series Forecasting, Nelson Siegel Model, Yield Curve.

JEL Classification: C53, F31, G17, C45.

Öz - Makine Öğrenmesi ve Derin Öğrenme Yaklaşımlarının Döviz Kuru Tahmini Konusundaki Tahmin Yeteneği

Bu çalışma, USD/TRY döviz kuru dalgalanmalarının tahmininde öngörü modellerinin etkinliğini değerlendirmektedir. Destek Vektör Makineleri (SVM), XGBoost, Uzun Kısa Süreli Bellek (LSTM) ve Kapalı Tekrarlayan Birim (GRU) modelleri 96 ve 21 özellik setiyle değerlendirilmiştir. Veriler 01.01.2010'dan 30.04.2024'e kadar Bloomberg, TCMB ve BDDK'dan alınmıştır. Bulgular, LSTM ve GRU modellerinin geleneksel modellerden daha üstün olduğunu göstermektedir. GRU en yüksek öngörü doğruluğunu sergilerken, SVM yüksek boyutlu verilerle kötü performans göstermekte, XGBoost ise orta düzeyde tahmin gücü sunmaktadır. Bu çalışma, finansal zaman serisi tahmininde model ve özellik seçiminin önemini vurgulamakta ve gelişmiş sinir ağlarının avantajlarını ortaya koymaktadır. Sonuçlar, analistler ve politika yapıcılar için sağlam ekonomik tahmin modelleri geliştirmede değerli bilgiler sunmaktadır.

Anahtar Kelimeler: Döviz Kuru, Makine Öğrenmesi, Derin Öğrenme, Zaman Serisi Tahmini, Nelson Siegel Modeli, Getiri Eğrisi

JEL Sınıflandırması: C53, F31, G17, C45.

---

* Corresponding Author, Banking Regulation and Supervision Agency - E-mail: fturkoglu@bddk.org.tr
ORCID: 0009-0007-4210-5408.
** Banking Regulation and Supervision Agency - E-mail: egocecek@bddk.org.tr - ORCID: 0009-0004-1117-6578.
*** Dr., Banking Regulation and Supervision Agency - E-mail: yyumrukuz@bddk.org.tr - ORCID: 0009-0002-3588-331X.

Article Received Date: 25.07.2024 Article Accepted Date: 11.09.2024

Cite: Türkoğlu, F., Göçecek, E. ve Yumrukuz Y. (2024). Predictive Abilities of Machine Learning and Deep Learning Approaches for Exchange Rate Prediction. *Journal of BRSA Banking and Financial Markets,* 18(2), 186-210. http://doi.org/10.46520/bddkdergisi.1600294.

1. Introduction

Structural and long-term factors establish a framework for short-term fluctuations in exchange rates. To comprehend the short-term movements of exchange rates, it is essential to consider factors such as changes in monetary policy, production levels, and differences in interest and inflation rates. These elements dictate how exchange rates oscillate within a certain range. For an extended period, yield curve parameters — market information and sentiments embedded within term structures — have been integral in explaining the dynamics behind exchange rates. According to previous studies, there is a potential close relationship between macro-financial variables and yield curve parameters. Consequently, yield curve parameters are regarded as robust proxies for macro-financial variables and are utilized in analyzing exchange rates, either as substitutes or in conjunction with macro-financial factors, provided the included factor is assumed to be mutually independent of yield curve parameters. Therefore, any model analyzing the relationship between interest rate differentials and exchange rate changes over different terms from 1 to 12 months should be specified within the framework developed by Chen and Tsang (2013) and subsequent studies. However, this framework must be modified and customized to account for the effects of currency substitution on the Turkish Lira (TL) demand function, the use of multiple monetary policy tools to implicitly manage the exchange rate, and the constraints imposed by FX (Foreign Exchange) reserves and the current account balance. Forecasting macroeconomic and financial time series is considered one of the most challenging applications of modern time series forecasting due to the inherently noisy, non-stationary, and chaotic nature of these series (G. Deboeck, 1994). This challenge is particularly pronounced when designing a model that includes periods of stress, as seen in Turkey or similar countries. Factors such as high volatility due to currency substitution, persistent current account deficits, and the use of multiple monetary policy tools must be taken into account. When forecasting such time series data, a common assumption is that the past behavior of the series contains all the information needed to predict its future behavior. Therefore, traditional attempts to forecast exchange rates have primarily focused on univariate time series analysis using approaches like Auto-Regressive Integrated Moving Average (ARIMA) and Random Walk (RW). The main issue with ARIMA is that it is a univariate model based on two key assumptions: the time series being forecasted is linear and stationary. Moreover, univariate time series models do not account for the effects of other parameters that might be crucial in determining the future value of a specific macroeconomic variable (Nyoni, Thabani, 2018). This paper aims to focus on multivariate time series forecasting. Additionally, in this paper, we study Support Vector Machines (SVM), XGBoost, and Neural Network models. Neural networks are known to be much more effective in mapping the dynamics of non-stationary time series due to their unique non-parametric, non-assumable, noise-tolerant, and adaptive properties. Neural networks are well-known function approximators that can map any nonlinear function without any prior assumptions about the data. However, Multilayer Perceptron (MLP) models based on Artificial Neural Networks (ANNs) often face problems such as overfitting, error degradation during backpropagation, and the inability to automatically determine optimal time delays when fitting time series data (Kamruzzaman, 2003). Therefore, we propose a Recurrent Neural Network using Long Short-Term Memory (LSTM), which can more effectively capture the nonlinearity and randomness of time series data, as well as overcome the problem of error corruption backpropagating through memory blocks. This shows superior capabilities for time series prediction with long temporal dependence. With the ability to memorize long historical data and automatically determine optimal time delays, LSTM and GRU RNN achieves higher prediction accuracy and generalizes well across different prediction intervals.

2. Literature Review

Balance of payments (BOP) consists of two main components representing real and financial flows in an open economy: the current account and the financial account. When capital mobility is perfect, financial markets face no constraints, allowing for the free movement of financial capital across borders. Similarly, in the absence of trade barriers, the real flow of goods and services across borders encounters no restrictions. The relationship between interest rates and exchange rates can be directly constructed based on financial flows. According to the Uncovered Interest Rate Parity (UIRP) (Fama, 1984), if there is a gap between domestic and foreign interest rates over a given term, the domestic currency must appreciate or depreciate accordingly. The expected depreciation (or appreciation) is crucial when analyzing exchange rates based on UIRP. The flow of money in financial markets is generally rapid, while cross-border flows from commercial activities are relatively slower. A country consistently running a current account deficit faces pressure for its currency to depreciate. Therefore, trade balance or current account balance is essential for assessing the fair level of the exchange rate, particularly in the medium to long term. Mathematically, UIRP can be expressed as:

$$y_t = \beta_0 + \sum_{j=1}^{p} \beta_j y_{t-j} + \epsilon_t \; (1)$$

where id is the domestic interest rate, if is the foreign interest rate, $E(St+1)$ is the expected future spot exchange rate, and St is the current spot exchange rate.

In practice, a trade deficit cannot be financed indefinitely by ongoing capital inflows because debt repayment requires achieving a current account surplus, even if asset sales or other extraordinary income sources might delay the point at which the position becomes unsustainable. Thus, relative interest rates combined with the trade balance determine the exchange rate level. When we look at the models we use in our article; Support Vector Machines (SVM) offer a novel approach for multivariate prediction and generalization. SVMs adopt a Structural Risk Minimization (SRM) approach, aiming to minimize the upper bound of generalization error rather than just the training error. This results in better generalization compared to traditional techniques (Vapnik, 1995). SVM regression is non-parametric due to its use of kernel functions, which allow it to model complex, nonlinear relationships.

In ϵ-SVM regression (L1 loss), the objective is to find a function $f\ (x)$ that deviates from the actual values yi by no more than ϵ and is as flat as possible. This is achieved by solving a Lagrangian function subject to the Karush-Kuhn-Tucker (KKT) conditions (Karush, 1939; Kuhn & Tucker, 1951). The optimization problem can be formulated as:

$$\min_{w,b,\xi} \left(\frac{1}{2} |w|^2 + C \sum_{i=1}^{n} \xi_i \right) (2)$$

Subject to:

$$y_i - (w \cdot \phi(x_i) + b) \leq \epsilon + \xi_i \; (3)$$

$$(w \cdot \phi(x_i) + b) - y_i \leq \epsilon + \xi_i \; (4)$$

$$\xi_i \geq 0 \; (5)$$

where $\mathbf{w}$ is the weight vector, b is the bias, ξi are slack variables, C is the regularization parameter, and $\phi(xi)$ is the feature mapping function.

However, SVMs often perform poorly if the training dataset is large or very noisy. On the other hand, SVMs can model nonlinear relationships thanks to their kernel functions (e.g., radial basis function, polynomial kernel). Extreme Gradient Boosting (XGBoost) is another powerful machine learning algorithm that uses decision trees and learns sequentially by correcting the mistakes of previous trees. XGBoost offers superior performance on larger datasets and high-dimensional data compared to SVM. Its boosting technique, known for fast computation capabilities, continuously increases model accuracy (Chen & Guestrin, 2016). The objective function of XGBoost can be defined as:

$$\mathcal{L}(\theta) = \sum_{i=1}^{n} l\,(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \; (6)$$

where l is a differentiable convex loss function that measures the difference between the prediction yi and the target yi, and Ω is a regularization term to control the complexity of the model.

Long Short-Term Memory (LSTM) networks, a type of Recurrent Neural Network (RNN), model dependencies between consecutive data points in time series data. Unlike XGBoost, which operates on independent data points, LSTMs incorporate time dependencies into the model, making them suitable for sequential data (Hochreiter & Schmidhuber, 1997).

LSTM networks are designed to avoid the long-term dependency problem, with the ability to remember information for extended periods. Each LSTM unit consists of a cell, an input gate, an output gate, and a forget gate, which regulate the flow of information. The cell state Ct is updated as follows (Hochreiter & Schmidhuber, 1997):

$$C_t = f_t * C_{t-1} + i_t * \widetilde{C}_t \; (7)$$

where ft is the forget gate, it is the input gate, and Ct is the candidate cell state. The output ht of the LSTM unit is given by:

$$h_t = o_t * \tanh(C_t) \; (8)$$

where ot is the output gate. The gates are defined as:

$$f_t = \sigma\big(W_f \cdot [h_{t-1}, x_t] + b_f\big) \; (9)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \; (10)$$

$$\widetilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \; (1$$

where σ is the sigmoid activation function, tanh is the hyperbolic tangent function, and W and b are weights and biases.

LSTMs' ability to store and access long-term information helps mitigate the vanishing gradient problem, making them effective for time series analysis. On the other hand, Gated Recurrent Unit (GRU) networks, a type of Recurrent Neural Network (RNN), model dependencies between consecutive data points in time series data. Like LSTMs, GRUs incorporate time dependencies into the model, making them suitable for sequential data (Cho et al., 2014). GRUs are designed to simplify the architecture of LSTMs while maintaining similar performance and effectiveness in capturing long-term dependencies.

Each GRU unit consists of an update gate and a reset gate, which regulate the flow of information. The GRU updates the hidden state ht as follows:

$$h_t = (1 - z_t) \circ h_{t-1} + z_t \circ \widetilde{h}_t \; (13)$$

where zt is the update gate and ht is the candidate hidden state. The update gate zt determines how much of the previous hidden state needs to be updated with new information, and it is defined as:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \; (14)$$

The candidate hidden state ht is calculated using the reset gate rt, which determines how much of the previous hidden state to forget:

$$\widetilde{h}_t = \tanh(W_h \cdot [r_t \circ h_{t-1}, x_t] + b_h) \; (15)$$

The reset gate rt is defined as:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \ (16)$$

where σ is the sigmoid activation function, tanh is the hyperbolic tangent function, and W and b are weights and biases.

GRUs simplify the architecture of LSTMs by combining the forget and input gates into a single update gate and omitting the cell state, making them more computationally efficient. This simplification helps maintain performance while reducing computational complexity, making GRUs effective for time series analysis. A visual comparison of different sequence models is given below.

Figure 1: Recurrent Neural Network (RNN) Architecture

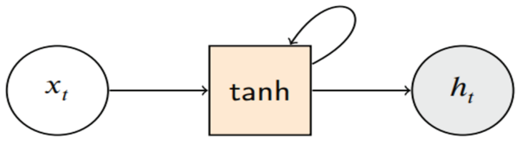


Figure 2: Long Short-Term Memory (LSTM) Network Architecture

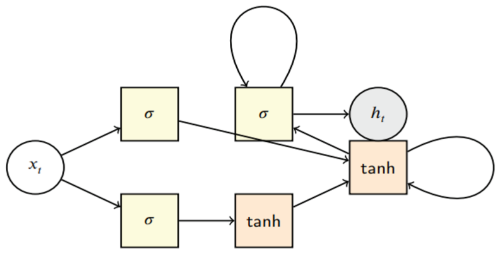
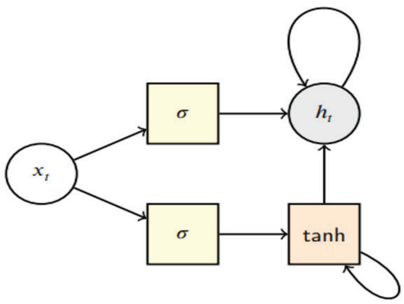


Figure 3: Gated Recurrent Unit (GRU) Network Architecture



Moreover, as the performance measurement metrics MSE and R-squared used in our article. R^2, also known as the coefficient of determination, measures the explanatory power of a regression model. It indicates how much of the total variance in the dependent variable can be explained by the independent variables. The value of R^2 ranges from 0 to 1, with values closer to 1 indicating that the model explains a larger portion of the variance in the dependent variable.

The Mean Squared Error (MSE) is a common measure used to evaluate the accuracy of a regression model. It calculates the average of the squares of the errors—that is, the average squared difference between the actual values and the predicted values.

The R^2 (R-squared) value, also known as the coefficient of determination, is a statistical measure that represents the proportion of the variance for the dependent variable that's explained by the independent variables in the model. It provides an indication of the goodness of fit and ranges from 0 to 1, where higher values indicate a better fit.

Below are the formulas for MSE and R^2:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (17)$$

where yi are the actual values, y^i are the predicted values, and n is the number of data points.

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}} \quad (18)$$

Where:

- *SS*res (Residual Sum of Squares): The sum of the squares of the residuals,
- *SS*tot (Total Sum of Squares): The total sum of squares, calculated as:

$$SS_{\text{tot}} = \sum_{i=1}^{n}(y_i - \bar{y})^2 \quad (19)$$

Here, yi are the actual values and y is the mean of the actual values.

$$SS_{\text{res}} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (20)$$

Mean Squared Error, measures the average squared difference between the actual and predicted values. It evaluates the performance of a model by calculating the mean of the squares of the errors. A smaller MSE indicates a better-performing model.

In conclusion, understanding the balance of payments and exchange rate dynamics, along with the application of advanced machine learning techniques like SVM, XGBoost, LSTM and GRU, provides a comprehensive framework for financial analysis and prediction. These methodologies, when combined, offer powerful tools for modeling complex economic relationships and forecasting future trends.

The study by Kaushik and Giri (2020) compares the performance of three different approaches for forecasting the USD/INR exchange rate: Vector Auto Regression (VAR), Support Vector Machine (SVM), and Long Short-Term Memory (LSTM). The analysis uses monthly historical data from April 1994 to December 2018 and incorporates several macroeconomic variables such as CPI, interest rates, and money supply for both the U.S. and India. The authors found that traditional econometric models like VAR are outperformed by contemporary machine learning techniques. Specifically, the LSTM model, a deep learning approach, achieved the highest accuracy at 97.83%, followed by SVM with 97.17%, while VAR showed a lower accuracy of 96.31%. The results suggest that modern methods, particularly LSTM, are better suited for capturing non-linear relationships in time series data compared to traditional econometric models.

In the study by Goncu (2019), machine learning techniques such as Ridge regression, decision tree regression, support vector regression (SVR), and linear regression are used to predict the USD/TRY exchange rate. The model incorporates macroeconomic variables including the domestic money supply, real interest rates, and the U.S. Federal Funds rate. The dataset consists of 148 monthly observations from January 2007 to May 2019. Among the techniques tested, Ridge regression outperforms the others with the lowest prediction errors and standard deviation, making it a useful tool for investors and policymakers. The study highlights the effectiveness of machine learning algorithms in capturing the relationship between macroeconomic factors and exchange rates, providing accurate predictions with potential policy implications, such as analyzing the impact of interest rate changes on exchange rates.

In the study by Ajumi and Kaushik (2018), the authors provide a literature review on the application of machine learning and deep learning methods for exchange rate forecasting. They compare traditional time series models like ARIMA and GARCH with more advanced techniques such as Support Vector Machines (SVM), Neural Networks, and hybrid models. The review emphasizes the limitations of traditional models in capturing non-linear patterns in financial data, while highlighting the effectiveness of deep learning methods in improving prediction accuracy. The study concludes that machine learning models, particularly hybrid approaches combining different algorithms, tend to outperform traditional econometric models in exchange rate forecasting. The review also notes the increasing application of these methods in emerging economies, where exchange rate volatility is a major issue. However, the authors caution that no single forecasting model is consistently superior across different datasets and market conditions.

In the study by Dautel et al. (2020), the authors explore the application of deep recurrent neural networks (RNNs) in foreign exchange (Forex) rate forecasting, specifically comparing Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) architectures. The models are evaluated against simpler neural network structures such as feedforward neural networks (FNNs) and traditional RNNs. The findings suggest that while deep learning models, particularly LSTMs and GRUs, outperform simpler models in terms of directional accuracy, they do not always yield better profitability in trading strategies. Surprisingly, simpler architectures like FNNs can sometimes achieve similar or even superior trading performance. The study emphasizes the challenge of balancing statistical performance with practical economic outcomes in exchange rate forecasting.

In their 2020 study, Kim et al. investigate the predictive performance of the Nelson-Siegel model alongside several machine learning techniques—Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Support Vector Regression (SVR), and Group Method of Data Handling (GMDH)—for forecasting the term structure of credit default swaps (CDS). Using CDS data from 2008 to 2019, the study analyzes the predictability of sovereign CDS spreads. The authors find that machine learning models consistently outperform the Nelson-Siegel model, with GMDH showing the highest prediction accuracy across multiple periods. Additionally, the study highlights the importance of distinguishing between volatile and stable periods when forecasting CDS term structures, noting that models performed worse during volatile periods, such as the 2008 financial crisis. This research underscores the value of data-driven methods over traditional model-driven approaches in predicting the future risk associated with credit defaults.

3. Data and Methodology

In the aforementioned literature, it is possible to conduct an analysis using daily or monthly datasets. We opted to work with daily data in our study. Our dataset comprises daily data from 2010 to 2024, sourced predominantly from Bloomberg, the Central Bank of the Republic of Turkey (CBRT), and the Banking Regulation and Supervision Agency (BDDK). Initially, the zero-coupon bond yields were calculated using the Fama-Bliss approach, followed by the estimation of the yield curve parameters using the Nelson-Siegel Model (NSM).

Both the Fama-Bliss approach for calculating zero-coupon bond yields and the NSM for estimating yield curve parameters are widely recognized methods in the literature. For example, Kim, W. J., Jung, G., & Choi, S.-Y. (2020) utilized a similar approach, combining the NSM with machine learning techniques to forecast CDS term structures. We chose these methods because the NSM is a well-established technique for parametric modeling of yield curves, and the Fama-Bliss method is a reliable approach for calculating zero-coupon bond yields. Furthermore, our use of daily data allows us to assess the model's performance with high-frequency data, aligning with our objective to generate more precise and up-to-date forecasts.

The primary objective of our study is to capture the fluctuations of the Turkish Lira (TRY) against the US Dollar (USD) in out-of-sample data with a one-month forecast horizon. The target variable is the monthly change in the USD/TRY exchange rate one day ahead. Our dataset includes primarily 20 variables, updated daily, categorized into three groups: 1) Turkey Macro Variables, 2) Turkey Financial

Variables, 3) the US Macro-financial Variables, and 4) Global Economic Indicators. Upon applying the Augmented Dickey-Fuller (ADF) test, we observed that some of our variables were non-stationary.

To address non-stationarity, we first applied the Box-Cox transformation to the non-stationary variables. Subsequently, we calculated the differences or percentage changes for certain variables. To enhance the predictive power regarding exchange rate movements, we generated new features by computing the differences or percentage changes over 15, 30, 45, 60, 75, and 90 days for all variables. In this study, feature engineering was employed to create lags of the raw data (15, 30, 45, 60, 75, and 90 days), and analyses were conducted both with and without these lags to compare results. The target is the monthly change in the exchange rate 30 days ahead.

One of the main challenges in this analysis is the high volatility of the Turkish Lira and the frequent market interventions in recent times, which have significantly impacted our prediction metrics and deteriorated our adjusted R-squared values.

The data utilized in this study were gathered from multiple authoritative sources: Bloomberg for financial and economic data, the CBRT for macroeconomic indicators and financial statistics, and the BRSA for banking sector data. The data preprocessing involved several steps to ensure the robustness of the analysis. Zero-Coupon Bond Yields Calculation, utilized the Fama-Bliss approach to derive zero-coupon bond yields. Yield Curve Estimation, applied the Nelson-Siegel model to estimate yield curve parameters. To address non-stationarity, we conducted the following transformations: Box-Cox Transformation, applied to non-stationary variables to stabilize variance and make the data more normally distributed; Differencing and percentage change calculations, performed differencing or calculated percentage changes for non-stationary variables and we managed to maket he data stationary (see Appendix A for detailed test results).

Additionally, we engineered new features by calculating the differences or percentage changes over various time windows (15, 30, 45, 60, 75, and 90 days) to capture temporal dynamics and enhance predictive accuracy. In this study, feature engineering was employed to create lags of the raw data (15, 30, 45, 60, 75, and 90 days), and analyses were conducted both with and without these lags to compare results. Furthermore, we dropped data with a correlation greater than 70% to avoid multicollinearity issues.

Our dataset comprises 5,114 daily observations from 2010 to 2024, containing key macro-financial variables across four categories: 1) Turkey Macro Variables, 2) Turkey Financial Variables, 3) US Macro-financial Variables, and 4) Global Economic Indicators. Descriptive statistics for these variables are provided in Appendix B, giving a detailed overview of the dataset's characteristics.

The descriptive analysis reveals that the dataset captures a wide range of financial market dynamics. For instance, the average CDS (credit default swap) spread during the sample period is 0.0003 with a skewness of 0.57, indicating slight asymmetry in the distribution. However, more interestingly, the XAU (gold prices) series exhibits extreme skewness (-6.09) and kurtosis (391.30), revealing the presence of significant outliers and reflecting the high volatility in gold prices during certain periods of the dataset. The maximum value of XAU reaches 928.78, which underscores sharp spikes in gold prices, particularly during global financial uncertainties.

Similarly, the VIX index, a common proxy for market volatility, has a mean value of 18.36 and a standard deviation of 7.03, reflecting the frequency of market stress events during the period, such as the European debt crisis and geopolitical tensions affecting Turkey. The VIX also shows positive skewness (2.27), indicating that market volatility spikes tend to be more pronounced than periods of relative calm. The BIST100 index, on the other hand, demonstrates relatively stable behavior with a near-zero mean and moderate fluctuations, as indicated by its small standard deviation and low skewness (-0.34).

The descriptive analysis of these variables highlights the complex and diverse nature of the dataset, which includes both stable indicators and highly volatile ones, making it a challenging dataset for predictive modeling. This variation across the dataset is crucial for testing the robustness and generalizability of the models.

We implemented four predictive models to forecast the monthly change in the USD/TRY exchange rate: Support Vector Machine (SVM), XGBoost, LSTM Networks, and GRU. SVM is a supervised learning algorithm used for classification and regression analysis, known for its robustness but potential limitations with high-dimensional data. XGBoost is a powerful ensemble learning method that excels in handling large datasets and is known for its computational efficiency and ability to capture complex patterns through decision trees. LSTM and GRU are types of recurrent neural networks (RNNs) designed to learn long-term dependencies in sequential data, particularly useful for time-series forecasting tasks like exchange rate prediction.

The distribution of variables, provided in Appendix C, reveals important insights into the underlying characteristics of our dataset. Many of the variables, including CDS spreads, VIX, and BIST100, follow distributions that are approximately normal with moderate skewness. However, some variables, such as XAU (gold prices), exhibit significant skewness (-6.09) and extreme kurtosis (391.30), indicating the presence of substantial outliers and periods of extreme volatility. These findings reflect the nature of global financial markets during periods of uncertainty, which significantly impacted Turkey's economy and financial indicators. The high volatility in gold prices, for instance, may be linked to global events such as the Eurozone crisis and geopolitical conflicts.

The outlier analysis in Appendix D further supports these observations, revealing extreme values in variables like XAU and the TRY1W Repo rate. These outliers highlight periods of sharp financial intervention or policy changes within the Turkish financial system. While these outliers provide valuable information about market stress and fluctuations, they also introduce challenges for model stability and accuracy. Spikes in variables such as the TRY1W Repo rate, in particular, reflect aggressive monetary policy actions by the Turkish Central Bank aimed at stabilizing the currency, which is critical for understanding broader fluctuations in the USD/TRY exchange rate.

The combination of diverse distributions and the presence of outliers across various macroeconomic and financial variables adds complexity to the predictive modeling process. These characteristics, ranging from moderate fluctuations to sharp market interventions, must be carefully accounted for in the modeling approach to ensure robust forecasting, particularly in periods of heightened market instability.

The models were developed using a time series analysis approach appropriate for our dataset, which comprises 5,114 observations. Instead of employing a sliding window approach, we divided the dataset into a training set comprising 95% of the data and a test set comprising the remaining 5%. This division, consistent with time series analysis, ensures that the data remains in chronological order, preserving the temporal structure of the series for accurate forecasting.

The training set, spanning from 2010 to a specific cutoff date, was used to train the models. The remaining data from the cutoff date to 2024 formed the test set. This method ensured that the models were validated on out-of-sample data, maintaining the temporal integrity of the time series.

At the end of the training phase, each model's performance was evaluated on the test set. This approach helped in capturing temporal dependencies and ensured that the models were robust across different time periods, providing reliable forecasts for the monthly change in the USD/TRY exchange rate.

The performance of these models was evaluated using MSE, and the predictive accuracy was assessed. The models were trained and validated on different subsets of the data to ensure robustness and generalizability. Despite the high volatility and frequent market interventions affecting the Turkish Lira, these models provided valuable insights into the exchange rate movements. One significant challenge encountered during the analysis was the high degree of market volatility, which introduces noise and reduces the predictability of exchange rate movements. In the context of Turkey, the exchange rate is particularly susceptible to both domestic and international political events, economic policy changes, and sudden shifts in investor sentiment. These factors contribute to the high volatility observed in the USD/TRY exchange rate.

Frequent interventions by the Central Bank and other regulatory bodies further complicate the modeling process by introducing sudden and unpredictable shifts in the data. Such interventions, aimed at stabilizing the currency, can cause abrupt changes that are difficult to forecast with traditional models. The unpredictability and rapid changes in market conditions negatively impacted our adjusted R-squared values, reflecting the difficulty in achieving high predictive accuracy under such volatile conditions.

Moreover, we observed that the predictive performance varied significantly across the different models. While the SVM model offered robust predictions for certain periods, it struggled during times of extreme volatility. The XGBoost model, with its ensemble learning approach, demonstrated better adaptability to the fluctuating market conditions but still faced challenges in capturing the abrupt changes caused by regulatory interventions. The LSTM model, designed to handle sequential data, provided valuable insights into long-term dependencies but was also affected by the high noise levels in the data. In this study, we performed hyperparameter tuning for three different models: Support Vector Regression (SVR), XGBoost, and GRU. Hyperparameter tuning is a crucial step in machine learning as it involves optimizing the parameters that directly influence the model's performance.

The goal is to find the best set of hyperparameters that yield the lowest MSE and highest R^2 score on the test dataset. For each model, we defined a grid of hyperparameters and used RandomizedSearchCV to perform the search. The best parameters were selected based on the performance metrics. Below, we present the hyperparameter values and the best hyperparameters found for each model.

The structural breaks and regime shifts in the exchange rate data, often driven by policy changes or external shocks, pose significant challenges for model stability and prediction accuracy. These breaks can lead to substantial model misspecification if not adequately accounted for, further complicating the forecasting task. In conclusion, the results of these models and their respective MSE values are discussed in the following sections. The study highlights the complexities involved in predicting exchange rate movements in a highly volatile and regulated market such as Turkey's and underscores the need for advanced modeling techniques and continuous adaptation to evolving market conditions. Understanding the limitations and challenges inherent in such an environment is crucial for improving forecasting models and developing more resilient economic strategies.

Table 1: Detailed Hyperparameter Values

Model	Hyperparams	Values
SVM	C	0.1, 1, 10, 100
	Gamma	0.1, 0.01, 0.001, 0.0001
	Epsilon	0.01, 0.1, 0.2, 0.5
	Kernel	Linear, RBF, Poly
XGBoost	n_estimators	5, 20, 500, 600
	Max depth	2, 3, 10
	Learning rate	0.0001, 0.001, 0.01, 0.1
LSTM	Units	200, 100, 50, 60, 20, 30, 40
	Optimizer	Rmsprop, Adam
	Batch size	16, 32, 40, 56, 128, 256
	Epochs	10, 20, 100, 200
	Dropout rate	0, 0.2, 0.4, 0.5
GRU	Units	200, 100, 50, 60, 20, 30, 40
	Optimizer	Rmsprop, Adam
	Batch size	16, 32, 40, 56, 128, 256
	Epochs	10, 20, 100, 200
	Dropout rate	0, 0.2, 0.4, 0.5

Table 2: Chosen Values

Model	Hyperparams	21 Features	96 Features
SVM	C	10	10
	Gamma	0.01	0.01
	Epsilon	0.01	0.01
	Kernel	RBF	RBF
XGBoost	n_estimators	500	500
	Max depth	3	3
	Learning rate	0.1	0.1
LSTM	Units	60	60
	Optimizer	Rmsprop	Rmsprop
	Batch size	32	32
	Epochs	200	200
	Dropout rate	0.5	0.5
GRU	Units	40	40
	Optimizer	Rmsprop	Rmsprop
	Batch size	40	40
	Epochs	200	200
	Dropout rate	0.2	0.2

4. Findings

The visual data in Appendix E showcases the interactions and trends of various financial indicators from 01.01.2010 to 30.04.2024, providing a comprehensive overview of the factors influencing the USD/TRY exchange rate. The Credit Default Swap (CDS) and Volatility Index (VIX) are pivotal measures of market risk and uncertainty. Spikes in these indices typically signal heightened market stress, often leading to depreciation of the TRY against USD. For instance, during periods of global economic turmoil or financial crises, both CDS and VIX spikes correlate with significant movements in the USD/TRY exchange rate, reflecting increased risk aversion among investors. This heightened risk aversion drives investors towards safer assets, such as the USD, thereby weakening the TRY.

The BIST100 index and MXEF (Emerging Markets Index) illustrate similar patterns, underscoring that the performance of the Turkish stock market is closely linked to broader emerging market trends. A declining BIST100 index often coincides with a weaker TRY as foreign investors withdraw their funds, demand TRY to get long position in FX-denominated assets, and seek safer investments. This capital flight can exacerbate depreciation pressures on the TRY. The correlation between these stock indices and the USD/TRY exchange rate highlights the sensitivity of TRY to shifts in investor sentiment towards emerging markets, further influenced by geopolitical and economic developments.

Gold (XAU) and Oil Prices provide insights into the commodity markets' influence on the USD/TRY exchange rate. Gold often acts as a safe-haven asset, with its price increasing during periods of market uncertainty, indirectly affecting the USD/TRY parity. As investors flock to gold in times of uncertainty, the corresponding increase in gold prices can signify broader market concerns, which often lead to a stronger USD and a weaker TRY. Conversely, oil price fluctuations have a direct impact on Turkey's economy, given its status as an oil importer. Volatile oil prices can lead to economic instability, influencing the exchange rate through impacts on inflation, trade balances, and economic growth.

The EUR/USD exchange rate and the Economic Policy Uncertainty Index (USEPU) are also crucial. Increased policy uncertainty in the US, as reflected in the USEPU, can lead to fluctuations in the EUR/USD rate, which in turn affects the USD/TRY exchange rate. A weaker USD relative to the EUR can lead to a stronger TRY, and vice versa, highlighting the interconnectedness of global currency markets.

This relationship underscores how economic policies and uncertainties in major economies ripple through to affect emerging market currencies like the TRY. Turkey's FX revenue derived from export and tourism comes mostly from European countries while import and other FX expenditures weighted in USD payments. Therefore, any depreciation of USD against EUR would probably strengthen BOP position of Turkey.

The TRY1W Repo and Overnight Interest Rate are direct indicators of Turkey's monetary policy stance. Changes in these rates influence short-term borrowing costs and liquidity. It is supposed that higher interest rates typically attract foreign investment, strengthening the TRY, while lower rates can lead to depreciation. The Central Bank's policies on these rates reflect its broader economic strategy and response to inflationary pressures, capital flows, and economic growth.

The US10Y Treasury Yield provides insights into the US interest rate environment. Rising US Treasury yields can lead to capital outflows from emerging markets like Turkey, resulting in a weaker TRY as investors seek higher returns in US assets. The relationship between US10Y and USD/TRY is critical for understanding the impact of US monetary policy on the Turkish Lira. Higher US yields signify stronger economic conditions in the US, prompting a shift in investor preference towards USD-denominated assets and away from emerging market currencies.

The Level, Slope, and Curvature of the yield curve further inform expectations of future interest rates and economic conditions. Changes in these metrics can signal shifts in economic growth and inflation expectations, which influence investor behavior and subsequently, the USD/TRY exchange rate. For example, a steepening yield curve might indicate expectations of higher future economic growth and inflation, which could lead to changes in capital flows and exchange rate dynamics.

The interplay of global risk measures, commodity prices, monetary policies, and market sentiment collectively shape the dynamics of the exchange rate. This comprehensive analysis offers valuable insights for financial analysts and policymakers, enabling them to understand the multifaceted factors driving exchange rate fluctuations and to devise informed strategies to navigate the complex financial landscape. The ability to decipher these relationships is crucial for making predictive assessments and managing the risks associated with currency volatility.

The features presented in the table above constitute our primary variables. In this study, we utilized a total of 21 main features, each carefully selected for their relevance and potential impact on the prediction of the USD/TRY exchange rate. In addition to these primary features, we generated an additional 75 derived features. These derived features were created through various transformations and combinations of the primary features to enhance the model's ability to capture complex patterns and interactions within the data.

Despite the extensive feature engineering, the effectiveness of these derived features on the performance of LSTM networks and Gated Recurrent Units (GRU) can be variable. LSTM and GRU models are inherently designed to capture temporal dependencies and long-term relationships in sequential data. Their architectures are specifically tailored to address issues like vanishing gradients and to remember information over long periods, which are crucial for time-series forecasting.

However, the inclusion of a large number of derived features might not always significantly improve the performance of LSTM and GRU models. In some cases, the models may already effectively capture the necessary temporal patterns using the original set of features. The added complexity from the derived features could lead to overfitting, especially if the derived features do not provide additional informative value beyond what the primary features offer.

Furthermore, LSTM and GRU models can be sensitive to the quality and relevance of the input features. If the derived features introduce noise or redundant information, this could negatively impact the model's ability to learn meaningful patterns. It is also possible that the models might inherently filter out less relevant features during training, thereby diminishing the impact of the additional derived features.

Therefore, while the derived features are intended to enhance the model's ability to understand complex relationships within the data, their actual impact on the performance of LSTM and GRU models may vary. It is essential to conduct thorough experimentation and validation to determine the true efficacy of these features in improving model predictions. In this study, one of the key explanations we sought to observe was precisely this variability in the impact of derived features on the performance of LSTM and GRU models.

Table 3: Comparison of LSTM and GRU Models

Comparison of LSTM and GRU (96 Features)		
Model	**Total Params**	**Trainable Params**
LSTM	103,061	102,661
GRU	40,541	40,261
Comparison of LSTM and GRU (21 Features)		
Model	**Total Params**	**Trainable Params**
LSTM	85,061	84,661
GRU	31,541	31,261

The table provide a detailed comparison of LSTM and GRU models when applied to datasets with 96 and 21 features. For both datasets, the LSTM models exhibit larger output shapes and significantly higher parameter counts compared to the GRU models.

In the case of the dataset with 96 features, the LSTM model has a total of 103,061 parameters, with 102,661 of these being trainable and 400 non-trainable. On the other hand, the GRU model for the same dataset has a total of 40,541 parameters, with 40,261 trainable and 280 non-trainable. This highlights that the LSTM model is more complex and parameter-intensive than the GRU model for this feature set.

Similarly, for the dataset with 21 features, the LSTM model again shows a higher complexity with a total of 85,061 parameters (84,661 trainable and 400 non-trainable), whereas the GRU model has 31,541 parameters (31,261 trainable and 280 non-trainable).

These comparisons indicate the greater computational complexity and potential for capturing intricate patterns with LSTM models, albeit at the cost of increased computational resources compared to GRU models. The visual representations are provided below.

Figure 4: GRU Summary (21 Features)

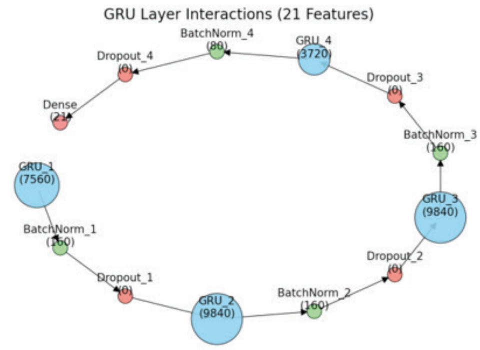


Figure 5: GRU Summary (96 Features)

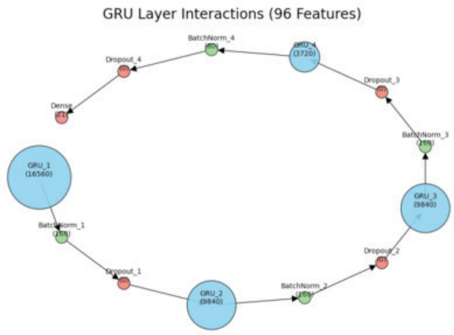


Figure 6: LSTM Summary (21 Features)

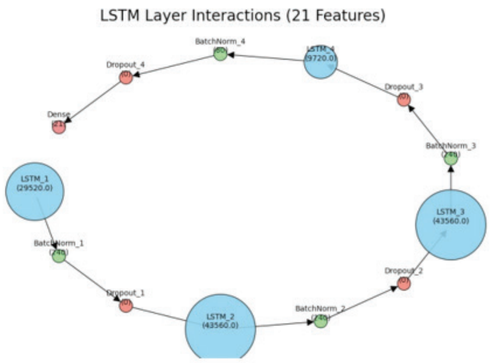
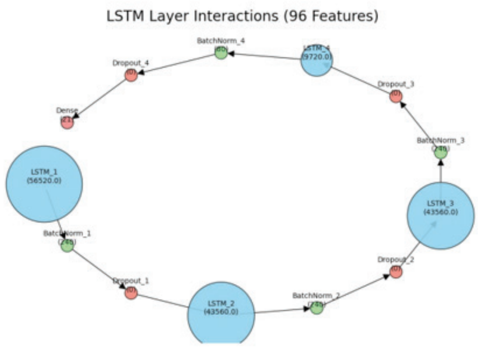


Figure 7: LSTM Summary (96 Features)



The comparative analysis of various forecasting models for exchange rate prediction highlights the efficacy of multivariate time series forecasting methods when dealing with noisy, non-stationary, and chaotic financial time series data. The study evaluates four models: SVM, XGBoost, LSTM, and GRU, using two different feature sets comprising 96 and 21 features, respectively.

Table 4: Model Results Comparison

Model Result (96 Features)

Model	MSE	R^2
SVM	0.000061428	-10.1373
XGBoost	0.000003212	0.4177
LSTM	0.000001639	0.7029
GRU	0.000001181	0.7859

Model Result (21 Features)

Model	MSE	R^2
SVM	0.000049790	-8.0514
XGBoost	0.000003674	0.3322
LSTM	0.000001005	0.8172
GRU	0.000000977	0.8224

For the feature set consisting of 96 features, the SVM model yields a MSE of 0.000061428 and an R^2 value of -10.1373. This negative R^2 value suggests that the SVM model performs poorly, failing to capture the underlying dynamics of exchange rate fluctuations, likely due to the complexity and high dimensionality of the feature set. The model's inability to effectively process the large number of features results in poor predictive performance, as evidenced by the high MSE and substantially negative R^2 value, indicating that the model's predictions are worse than a simple mean-based prediction.

The XGBoost model, on the other hand, performs significantly better with an MSE of 0.000003212 and an R^2 of 0.4177. This indicates moderate predictive power and a better handling of the feature set compared to SVM. XGBoost's ability to handle large feature sets through boosting and regularization techniques helps it capture the relevant patterns and dependencies in the data, leading to improved predictive accuracy. However, while the R^2 value is positive, it still suggests that there is room for improvement in capturing the full complexity of the exchange rate movements.

The LSTM model shows further improvement with an MSE of 0.000001639 and an R^2 of 0.7029, demonstrating its capability to capture long-term dependencies and nonlinear patterns in the data. LSTM's architecture, which includes memory cells to store information over long periods, allows it to effectively learn and model the temporal dependencies present in exchange rate time series data. This results in a substantial improvement in predictive performance compared to SVM and XGBoost, as indicated by the lower MSE and higher R^2 value.

The GRU model outperforms the other models with the lowest MSE of 0.000001181 and the highest R^2 of 0.7859. GRU's simpler architecture compared to LSTM, while still retaining the ability to capture long-term dependencies, enables it to achieve superior performance in time series prediction tasks. The model's ability to effectively process and learn from the feature set results in the highest predictive accuracy among the models evaluated, as evidenced by the lowest MSE and highest R^2 value.

When utilizing the reduced feature set of 21 features, the SVM model achieves an MSE of 0.000049790 and an R^2 of -8.0514. Although this indicates poor performance, it represents a slight improvement compared to its performance with the larger feature set. This suggests that SVM struggles with high-dimensional data but can achieve marginally better results with a reduced number of features. However, the negative R^2 value still indicates that the model's predictions are not reliable.

The XGBoost model shows an MSE of 0.000003674 and an R^2 of 0.3322, which, while slightly lower than its performance with 96 features, still indicates reasonable predictive power. XGBoost's robustness to overfitting and its ability to handle smaller feature sets effectively contribute to its relatively good performance. However, the decrease in R^2 value compared to the larger feature set suggests that some important information may be lost when reducing the number of features.

The LSTM model excels with an MSE of 0.000001005 and an R^2 of 0.8172, further solidifying its ability to model complex time series data effectively. LSTM's performance improvement with the reduced feature set indicates its strength in capturing the essential patterns and dependencies in the data without relying on a large number of features. This results in higher predictive accuracy, as evidenced by the lower MSE and higher R^2 value compared to the larger feature set.

The GRU model continues to demonstrate superior performance with an MSE of 0.000000977 and an R^2 of 0.8224, confirming its robustness and accuracy in forecasting tasks. GRU's ability to maintain high performance with a reduced number of features highlights its efficiency and effectiveness in processing and learning from the data. The lowest MSE and highest R^2 value among the models evaluated indicate that GRU is the most accurate and reliable model for predicting exchange rates in this study.

These results underscore the importance of selecting appropriate models and feature sets for time series forecasting in financial contexts. Neural network models, particularly LSTM and GRU, exhibit superior performance due to their ability to capture nonlinearity and temporal dependencies inherent in exchange rate data. The findings align with the literature that emphasizes the strengths of recurrent neural networks in handling complex, non-stationary financial time series and their potential for enhancing predictive accuracy in economic and financial forecasting. This study demonstrates that while traditional models like SVM may struggle with high-dimensional data, advanced models like LSTM and GRU can effectively leverage both large and small feature sets to provide accurate and reliable predictions.

Figure 8: Result Summary (21 Features)

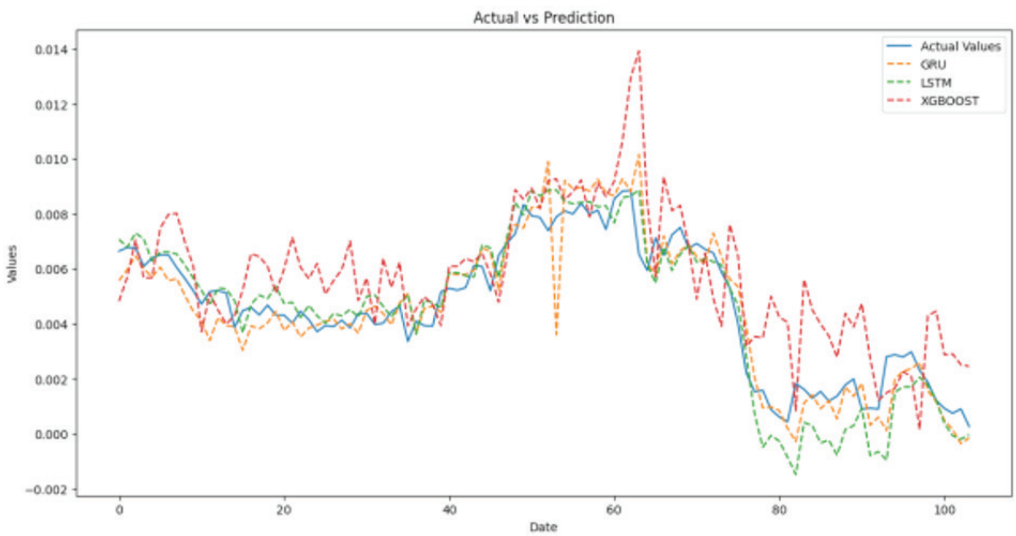
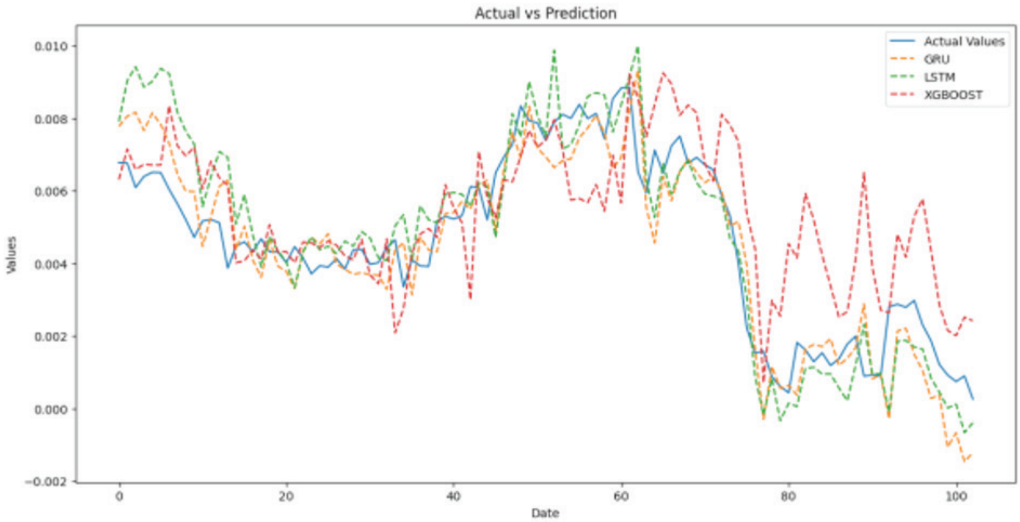


Figure 9: Result Summary (96 Features)



The two figures provided compare the actual values against the predictions made by GRU, LSTM, and XGBoost models for two different feature sets (96 features and 21 features).

In the first figure, representing the predictions with 96 features, the actual values (blue solid line) are plotted against the predictions made by GRU (green dashed line), LSTM (orange dashed line), and XGBoost (red dashed line). The GRU model closely follows the actual values for most of the time series, capturing the trends and fluctuations reasonably well, especially in periods of significant changes. The LSTM model also performs well, mirroring the actual values with slight deviations, indicating its effectiveness in modeling temporal dependencies. The XGBoost model, while capturing the general direction of the trends, shows more pronounced deviations and a higher level of noise compared to the neural network models, reflecting its lower performance as indicated by the quantitative metrics in the table.

In the second figure, which shows the predictions with 21 features, the GRU model again demonstrates a strong predictive capability, closely aligning with the actual values and effectively capturing both minor and major fluctuations. The LSTM model continues to perform well, with predictions that generally follow the actual values, though with occasional deviations that suggest it may struggle slightly more with the reduced feature set. The XGBoost model shows more significant deviations from the actual values, with a higher level of volatility and less accurate trend capture, indicating that it is less effective in handling the reduced feature set compared to the neural network models.

Overall, both figures illustrate the superior performance of the GRU and LSTM models in capturing the dynamics of exchange rate movements, particularly when using a larger feature set. The XGBoost model, while still useful, exhibits more noise and less accuracy in its predictions, highlighting the advantages of neural network models for this type of time series forecasting.

5. Conclusion

This study provides a comprehensive analysis of the predictive ability of various forecasting models for the USD/TRY exchange rate, utilizing a dataset comprising multiple macro-financial series. It addresses the challenges posed by the noisy, non-stationary, and chaotic nature of financial time series data. By evaluating four models—Support Vector Machine (SVM), XGBoost, Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU)—using two distinct feature sets (96 and 21 features), we gain significant insights into their performance and suitability for exchange rate forecasting.

Our findings demonstrate that neural network models, particularly LSTM and GRU, offer superior performance in capturing the intricate dynamics of exchange rate movements. The GRU model, with its simpler architecture and fewer parameters compared to LSTM, consistently outperforms the other models, achieving the lowest Mean Squared Error (MSE) and the highest R^2 values. This emphasizes GRU's efficiency and effectiveness in handling time series prediction tasks, especially in the context of financial data with long-term dependencies and nonlinearity.

In contrast, the SVM model performs poorly, particularly with the higher-dimensional feature set, reflecting its limitations in processing complex and high-dimensional data. The XGBoost model, while providing moderate predictive power, struggles to match the accuracy and reliability of the neural network models. This suggests that traditional machine learning approaches may not fully capture the temporal dependencies and nonlinearity inherent in exchange rate data.

The study also underscores the importance of feature engineering and careful selection of input variables. The models' performance varies depending on the number and type of features used, with the reduced feature set (21 features) in some cases improving results by reducing overfitting and focusing on the most relevant predictors. However, the robustness of the neural network models, particularly LSTM and GRU, across both feature sets highlights their flexibility and adaptability in time series forecasting.

In line with existing literature on the strengths of recurrent neural networks in handling complex, non-stationary financial data, our results suggest that neural networks, especially GRU and LSTM, are particularly effective for exchange rate forecasting. These insights offer valuable implications for financial analysts and policymakers aiming to develop more accurate and reliable forecasting models in volatile and dynamic market environments.

Future studies could explore hybrid models that combine neural networks with other machine learning techniques to further enhance predictive accuracy. Additionally, incorporating alternative macroeconomic indicators or experimenting with different temporal granularities (e.g., weekly or quarterly data) could yield further improvements in forecasting exchange rates.

6. Acknowledgments

We thank Dr. Levent Güntay, Dr. Emrah Ahi and Dr. Sedat Özer for his contribution.

Appendix

A)

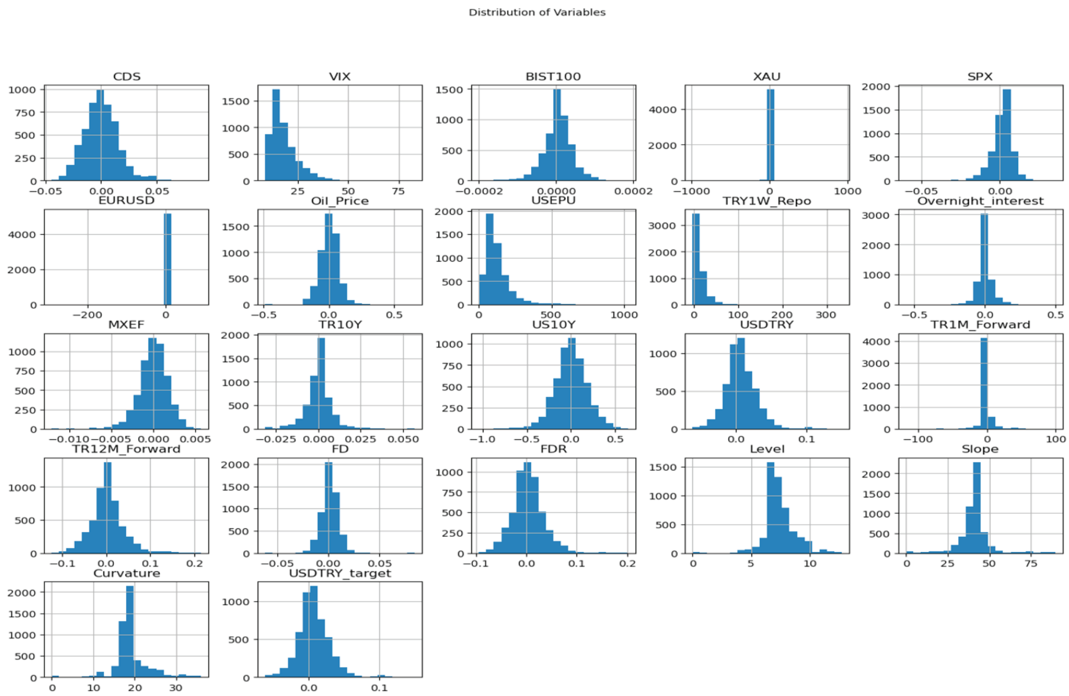
Series	Test Statistic	p-value	(1%)	(5%)	(10%)	Result
CDS	-78.139	0.0	-34.316	-28.621	-25.671	Stationary
VIX	-113.641	0.0	-34.316	-28.621	-25.671	Stationary
BIST100	-89.486	0.0	-34.316	-28.621	-25.671	Stationary
XAU	-78.085	0.0	-34.316	-28.621	-25.671	Stationary
SPX	-90.233	0.0	-34.316	-28.621	-25.671	Stationary
EURUSD	-149.871	0.0	-34.316	-28.621	-25.671	Stationary
Oil_price	-68.441	0.0	-34.316	-28.621	-25.671	Stationary
US_EPUI	-152.067	0.0	-34.316	-28.621	-25.671	Stationary
TRY1W_Rep	-120.014	0.0	-34.316	-28.621	-25.671	Stationary
Overnight_Interests	-107.418	0.0	-34.316	-28.621	-25.671	Stationary
MXEF	-79.238	0.0	-34.316	-28.621	-25.671	Stationary
TR10Y	-75.333	0.0	-34.316	-28.621	-25.671	Stationary
US10Y	-73.343	0.0	-34.316	-28.621	-25.671	Stationary
USDTRY	-78.223	0.0	-34.316	-28.621	-25.671	Stationary
TRY1M_forward	-106.209	0.0	-34.316	-28.621	-25.671	Stationary
TRY12M_forward	-75.519	0.0	-34.316	-28.621	-25.671	Stationary
FD	-80.343	0.0	-34.316	-28.621	-25.671	Stationary
FDR	-67.825	0.0	-34.316	-28.621	-25.671	Stationary
Level	-94.147	0.0	-34.316	-28.621	-25.671	Stationary
Slope	-88.769	0.0	-34.316	-28.621	-25.671	Stationary
Curvature	-98.988	0.0	-34.316	-28.621	-25.671	Stationary

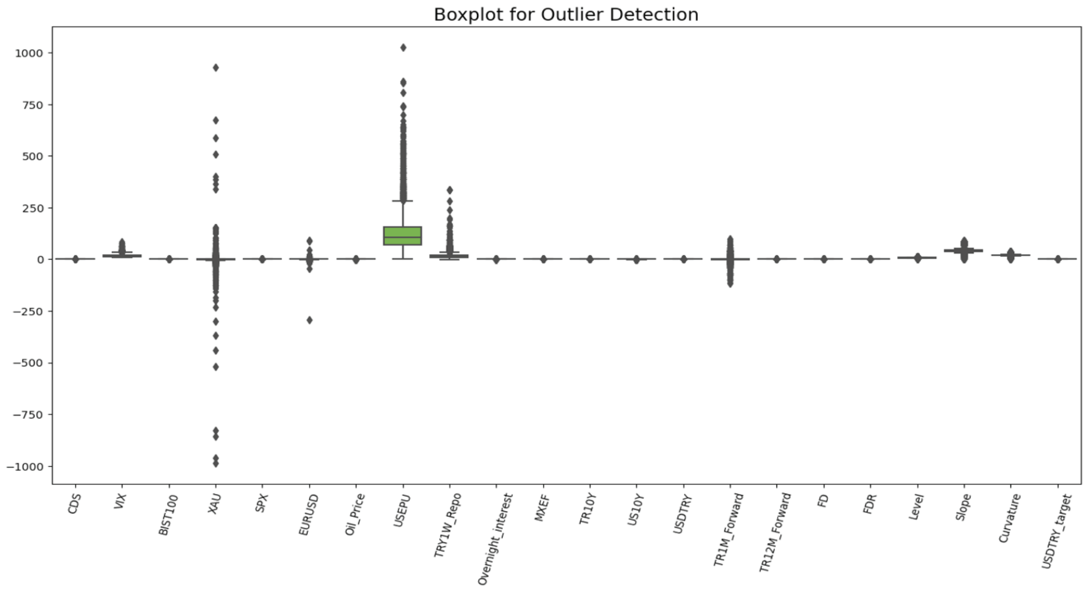
B)

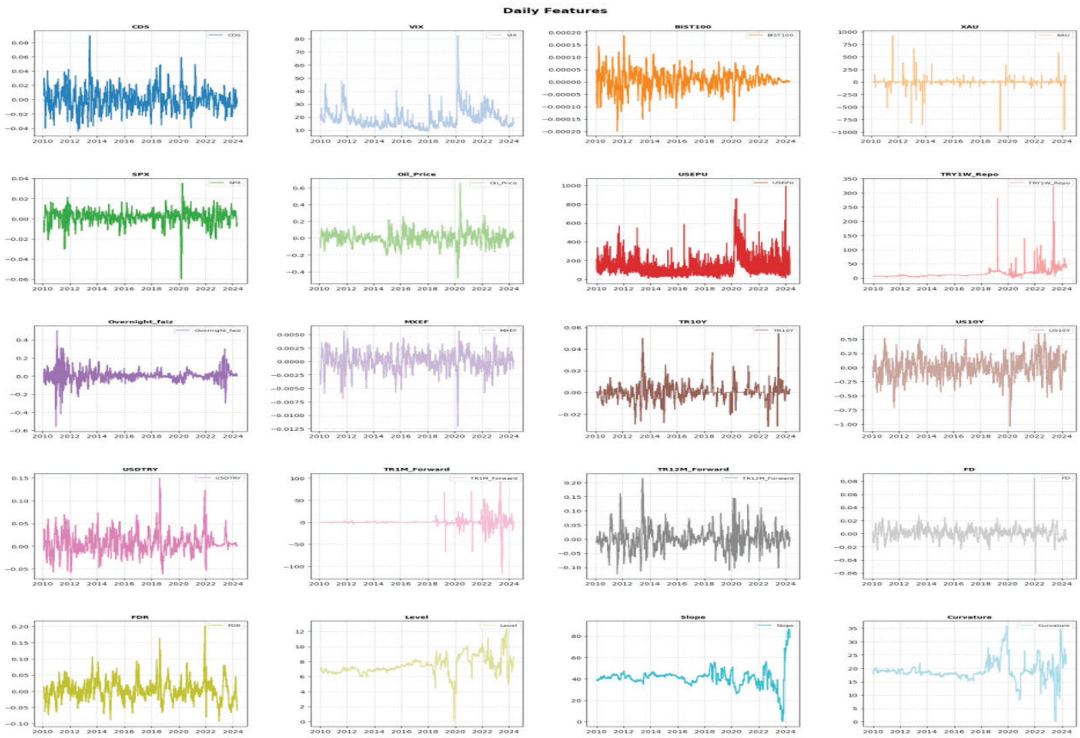
Series	Count	Mean	Std	Min	0,25	0,5	0,75	Max	Skewness	Kurtosis
CDS	5173	0.000291	0.015435	-0.044382	-0.009865	-0.00023	0.00958	0.089869	0.572925	1469734
VIX	5173	18362842	7030948	41883	13.54	16.52	45343	82.69	2269701	972615
BIST100	5173	0,000005	3.7e-05	-0.0002	-1.4e-05	0,000005	2.6e-05	0.000187	-0.33957	2348942
XAU	5173	-1157135	37822001	-986604746	-1925634	-1001785	0.126183	928785146	-6086674	391301459
SPX	5173	0.001432	0.006865	-0.059757	-0.001577	0.002394	0.005464	0.035439	-1586848	922129
EURUSD	5173	-0.040036	4606357	-291764147	-0.078629	-0.009798	0.060463	91134565	-46294245	3170151727
Oil_Price	5173	0.00395	0.078152	-0.485869	-0.039342	0.008562	0.046968	0.656402	-0.104576	7613573
USEPU	5173	1289575	91162995	11749	71.52	105.79	156.64	1026.38	2582198	10649047
TRY1W_Repo	5173	15690256	17333526	-3011	7594	10815	18186	334316	7424393	97511865
Overnight_Interest	5173	0.003796	0.053444	-0.556304	-0.017488	0.001231	0.022739	0.502198	0.171386	10017943
MXEF	5173	2.3e-05	0.00174	-0.012094	-0.000922	0.00013	0.001137	0.005722	-0.846406	3158803
TR10Y	5173	0.000336	0.008378	-0.03137	-0.002693	0.0	0.003462	0.057046	0.947493	8241225
US10Y	5173	0.00323	0.183285	-1041835	-0.106797	0.003789	0.115947	0.650589	-0.1808	1110548
USDTRY	5173	0.007479	0.022245	-0.061189	-0.005089	0.004738	0.019675	0.149435	0.886354	3383107
TR1M_Forward	5173	0.228788	10107271	-117516	-0.67	0.045	0.815	99853	-1146748	3618547
TR12M_Forward	5173	0.002577	0.037299	-0.123629	-0.017718	0.000807	0.019953	0.215293	0.796507	3034996
FD	5173	0.001528	0.007896	-0.061574	-0.003046	0.00168	0.006198	0.084493	0.026772	3410196
FDR	5173	0.004026	0.031685	-0.094013	-0.014971	0.002069	0.020938	0.202064	0.714995	2477681
Level	5173	7476998	127687	0.0	6811328	7295368	8047324	12755163	0.242961	4118932
Slope	5173	41169499	9242056	0.0	37927957	41797196	43805895	89623524	0.63534	8475983
Curvature	5173	19095559	4003703	0.0	17719113	18419531	19640179	35971433	0.493182	5771588
USDTRY_target	5173	0.007475	0.022245	-0.061189	-0.005089	0.004737	0.019667	0.149435	0.886876	3384193

B)

Series	Count	Mean	Std	Min	0,25	0,5	0,75	Max	Skewness	Kurtosis
CDS	5173	0.000291	0.015435	-0.044382	-0.009865	-0.00023	0.00958	0.089869	0.572925	1469734
VIX	5173	18362842	7030948	41883	13.54	16.52	45343	82.69	2269701	972615
BIST100	5173	0,000005	3.7e-05	-0.0002	-1.4e-05	0,000005	2.6e-05	0.000187	-0.33957	2348942
XAU	5173	-1157135	37822001	-986604746	-1925634	-1001785	0.126183	928785146	-6086674	391301459
SPX	5173	0.001432	0.006865	-0.059757	-0.001577	0.002394	0.005464	0.035439	-1586848	922129
EURUSD	5173	-0.040036	4606357	-291764147	-0.078629	-0.009798	0.060463	91134565	-46294245	3170151727
Oil_Price	5173	0.00395	0.078152	-0.485869	-0.039342	0.008562	0.046968	0.656402	-0.104576	7613573
USEPU	5173	1289575	91162995	11749	71.52	105.79	156.64	1026.38	2582198	10649047
TRY1W_Repo	5173	15690256	17333526	-3011	7594	10815	18186	334316	7424393	97511865
Overnight_Interest	5173	0.003796	0.053444	-0.556304	-0.017488	0.001231	0.022739	0.502198	0.171386	10017943
MXEF	5173	2.3e-05	0.00174	-0.012094	-0.000922	0.00013	0.001137	0.005722	-0.846406	3158803
TR10Y	5173	0.000336	0.008378	-0.03137	-0.002693	0.0	0.003462	0.057046	0.947493	8241225
US10Y	5173	0.00323	0.183285	-1041835	-0.106797	0.003789	0.115947	0.650589	-0.1808	1110548
USDTRY	5173	0.007479	0.022245	-0.061189	-0.005089	0.004738	0.019675	0.149435	0.886354	3383107
TR1M_Forward	5173	0.228788	10107271	-117516	-0.67	0.045	0.815	99853	-1146748	3618547
TR12M_Forward	5173	0.002577	0.037299	-0.123629	-0.017718	0.000807	0.019953	0.215293	0.796507	3034996
FD	5173	0.001528	0.007896	-0.061574	-0.003046	0.00168	0.006198	0.084493	0.026772	3410196
FDR	5173	0.004026	0.031685	-0.094013	-0.014971	0.002069	0.020938	0.202064	0.714995	2477681
Level	5173	7476998	127687	0.0	6811328	7295368	8047324	12755163	0.242961	4118932
Slope	5173	41169499	9242056	0.0	37927957	41797196	43805895	89623524	0.63534	8475983
Curvature	5173	19095559	4003703	0.0	17719113	18419531	19640179	35971433	0.493182	5771588
USDTRY_target	5173	0.007475	0.022245	-0.061189	-0.005089	0.004737	0.019667	0.149435	0.886876	3384193


Distribution of Variables
CDS
VIX
BIST100
XAU
SPX
EURUSD
Oil_Price
USEPU
TRY1W_Repo
Overnight_interest
MXEF
TR10Y
US10Y
USDTRY
TR1M_Forward
TR12M_Forward
FD
FDR
Level
Slope
Curvature
USDTRY_target


Boxplot for Outlier Detection
CDS
VIX
BIST100
XAU
SPX
EURUSD
Oil_Price
USEPU
TRY1W_Repo
Overnight_interest
MXEF
TR10Y
US10Y
USDTRY
TR1M_Forward
TR12M_Forward
FD
FDR
Level
Slope
Curvature
USDTRY_target


Daily Features
CDS
VIX
BIST100
XAU
SPX
Oil_Price
USEPU
TRY1W_Repo
Overnight_faiz
MXEF
TR10Y
US10Y
USDTRY
TR1M_Forward
TR12M_Forward
FD
FDR
Level
Slope
Curvature

References

1. Fama, E. F. (1984). Forward and spot exchange rates. *Journal of Monetary Economics*, 14(3), 319-338.
2. Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. J. Stat. Mech.Theory Exp. 2008, P10008.
3. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794).
4. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735-1780.
5. Kaushik, M., & Giri, A. K. (2020). *Forecasting Foreign Exchange Rate: A Multivariate Comparative Analysis between Traditional Econometric, Contemporary Machine Learning & Deep Learning Techniques*.
6. Goncu, A. (2019). *Prediction of Exchange Rates with Machine Learning*.
7. Ajumi, O., & Kaushik, A. (2018). *Exchange Rates Prediction via Deep Learning and Machine Learning: A Literature Survey on Currency Forecasting*.
8. Dautel, A. J., Härdle, W. K., Lessmann, S., & Seow, H.-V. (2020). *Forex Exchange Rate Forecasting Using Deep Recurrent Neural Networks*.
9. Kim, W. J., Jung, G., & Choi, S.-Y. (2020). *Forecasting CDS Term Structure Based on Nelson-Siegel Model and Machine Learning.* Complexity, 2020.
10. Chen, Q., Tsang, E., 2013. Forecasting macroeconomic and financial time series. Physica A, 392, 529–537.
11. Chen, Q., Wu, T.T., Fang, M., 2013. Detectinglocalcommunitystructure in complex networks based on local degree central nodes. Physica A. 392, 529–537.
12. Chen, T., Guestrin, C., 2016, XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785-794.
13. Cho, K., van Merriënboer, B., Bahdanau, D., Bengio, Y., 2014. On the properties of neural machine translation: Encoder-decoder approaches. Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation, 103-111.
14. Clauset, A., Newman, M.E.J., Moore, C., 2004. Finding community structure in very large networks. Phys. Rev. E. 70, 066111.
15. Danon, L., Diaz-Guilera, A., Duch, J., Arenas, A., 2005. Comparing community structure identification. J. Stat. Mech.-Theory Exp. , P09008.

16. Deboeck, G., 1994. Trading on the Edge: Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets. New York: John Wiley & Sons.
17. Fabio, D.R., Fabio, D., Carlo, P., 2013. Profiling core-periphery network structure by random walkers. Sci. Rep. 3, 1467.
18. Fabricio, B., Liang, Z., 2013. Fuzzy community structure detection by particle competition and cooperation. Soft Comput. 17, 659–673.
19. Fortunato, S., 2010. Community detection in graphs. Phys. Rep.-Rev. Sec. Phys. Lett. 486, 75–174.
20. Fortunato, S., Barthelemy, M., 2007. Resolution limit in community detection. Proc. Natl. Acad. Sci. U. S. A. 104, 36–41.
21. Gregory, S., 2011. Fuzzy overlapping communities in networks. J. Stat. Mech.-Theory Exp. , P02017.
22. Havens, T.C., Bezdek, J.C., Leckie, C., R.K., Palaniswami, M., 2013. A soft modularity function for detecting fuzzy communities in social networks. IEEE Trans. Fuzzy Syst. 21, 1170–1175.
23. Hochreiter, S., Schmidhuber, J., 1997. Long Short-Term Memory. Neural Computation, 9(8), 1735-1780.
24. Hullermeier, E., Rifqi, M., 2009. A fuzzy variant of the rand index for comparing clustering structures, in: in Proc. IFSA/EUSFLAT Conf., pp. 1294–1298.
25. Karush, W., 1939. Minima of functions of several variables with inequalities as side constraints. Master's Thesis, Department of Mathematics, University of Chicago.
26. Kamruzzaman, J., 2003. Multi-Layer Perceptron (MLP) for Time Series Prediction: A Review. Australian Journal of Intelligent Information Processing Systems, 8(2), 52-58.
27. Kuhn, H. W., Tucker, A. W., 1951. Nonlinear programming. Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 481-492.
28. Lancichinetti, A., Fortunato, S., 2009. Benchmarks for testing community detection algorithms on directed and weighted graphs with overlapping communities. Phys. Rev. E. 80, 016118.
29. Lancichinetti, A., Fortunato, S., Radicchi, F., 2008. Benchmark graphs for testing community detection algorithms. Phys. Rev. E. 78, 046110.
30. Li, J., Wang, X., Eustace, J., 2013. Detecting overlapping communities by seed community in weighted complex networks. Physica A. 392, 6125–6134.
31. Liu, J., 2010. Fuzzy modularity and fuzzy community structure in networks. Eur. Phys. J. B. 77, 547–557.
32. Liu, W., Pellegrini, M., Wang, X., 2014. Detecting communities based on network topology. Sci. Rep. 4, 5739.

33. Lou, H., Li, S., Zhao, Y., 2013. Detecting community structure using label propagation with weighted coherent neighborhood propinquity. Physica A. 392, 3095–3105.
34. Nepusz, T., Petróczi, A., Négyessy, L., Bazsó, F., 2008. Fuzzy communities and the concept of bridgeness in complex networks. Phys. Rev. E. 77, 016107.
35. Newman, M.E.J., Girvan, M., 2004. Finding and evaluating community structure in networks. Phys. Rev. E. 69, 026113.
36. Nyoni, Thabani., 2018. Univariate time series analysis and forecasting models in economics. Journal of Economics and Finance, 1(1), 1-15.
37. Psorakis, I., Roberts, S., Ebden, M., Sheldon, B., 2011. Overlapping community detection using bayesian non-negative matrix factorization. Phys. Rev. E. 83, 066114.
38. Raghavan, U., Albert, R., Kumara, S., 2007. Near linear time algorithm to detect community structures in large-scale networks. Phys. Rev E. 76, 036106.
39. Sobolevsky, S., Campari, R., 2014. General optimization technique for high-quality community detection in complex networks. Phys. Rev. E. 90, 012811.
40. Sun, P., Gao, L., Han, S., 2011. Identification of overlapping and non-overlapping community structure by fuzzy clustering in complex networks. Inf. Sci. 181, 1060–1071.
41. Vehlow, C., Reinhardt, T., Weiskopf, D., 2013. Visualizing fuzzy overlapping communities in networks. IEEE Trans. Vis. Comput. Graph. 19, 2486–2495.
42. Šubelj, L., Bajec, M., 2011a. Robust network community detection using balanced propagation. Eur. Phys. J. B. 81, 353–362.
43. Šubelj, L., Bajec, M., 2011b. Unfolding communities in large complex networks: Combining defensive and offensive label propagation for core extraction. Phys. Rev. E. 83, 036103.
44. Šubelj, L., Bajec, M., 2012. Ubiquitousness of link-density and linkpattern communities in real-world networks. Eur. Phys. J. B. 85, 1–11.
45. Vapnik, V., 1995. The Nature of Statistical Learning Theory. New York: Springer-Verlag.
46. Wang, W., Liu, D., Liu, X., Pan, L., 2013. Fuzzy overlapping community detection based on local random walk and multidimensional scaling. Physica A. 392, 6578–6586.
47. Wang, X., Li, J., 2013. Detecting communities by the core-vertex and intimate degree in complex networks. Physica A. 392, 2555–2563.
48. Zhang, S., Wang, R., Zhang, X., 2007. Identification of overlapping community structure in complex networks using fuzzy c-means clustering. Physica A. 374, 483–490.
49. Zhang, Y., Yeung, D., 2012. Overlapping community detection via bounded nonnegative matrix tri-factorization, in: In Proc. ACM SIGKDD Conf., pp. 606–614.