Journal of Money and Economy
Vol. 19, No. 2, Summer 2024
pp. **-**
DOI: 10.29252/jme.16.1.1

Original Research Article

Prediction the Short-term Exchange Rate of USD/IRR Using Deep Learning and the Impact of Sentiment Analysis Features on it

Seyed Jafar Mortazavian Farsani[*] **Abbas Toloie Eshlaghy**[†]
Reza Radfar[‡]

Received: 09 Apr 2025 Approved: 01 Jun 2025

This study investigates the role of sentiment analysis in improving exchange rate prediction models, providing empirical evidence for narrative economics; the idea that economic outcomes are shaped by prevailing beliefs and popular narratives. By integrating sentiment-based features into predictive frameworks, we demonstrate that exchange rate movements are influenced by subjective factors beyond traditional economic variables. Using data collected from Oct. 2019 to Mar. 2023, our findings suggest that market sentiment systematically impacts currency fluctuations. To assess the effectiveness of sentiment-enhanced models, we compare various forecasting approaches. Notably, a generalized linear model (GLM) outperforms more complex deep learning architectures, including long short-term memory (LSTM) networks and hybrid CNN-LSTM models. Additionally, even an optimized multilayer perceptron (MLP) fails to surpass GLM performance, highlighting the potential linearity of the relationship between predictors and exchange rates. These results underscore the importance of aligning model complexity with the statistical properties of the target variable. Beyond exchange rate forecasting, our study underscores the broader significance of incorporating sentiment and narratives into economic models. By acknowledging the role of subjective beliefs, researchers and policymakers can enhance predictive accuracy and improve decision-making processes in financial markets.

Keywords: Prediction, Narrative Economics, Foreign Exchange Rate, Deep Learning Models, Sentiment Analysis, Social Network
JEL Classification*:* A12, C45, C53, C51, C55, F31

[*]Department of Information Technology Management, Faculty of Management and Economics, Science and Research Branch, Islamic Azad University, Tehran, Iran; j.mortazavian@srbiau.ac.ir
[†]Department of Information Technology Management, Faculty of Management and Economics, Science and Research Branch, Islamic Azad University, Tehran, Iran (Corresponding Author); toloie@gmail.com
[‡] Department of Industrial Management, Faculty of Management and Economics, Science and Research Branch, Islamic Azad University, Tehran, Iran; r.radfar@srbiau.ac.ir

[DOI: 10.61882/jme.19.2.249]

1 Introduction

The exchange rate is a key factor in macroeconomics, influencing international trade and the economy of each country (Morina et al., 2020). According to numerous economic theories, it reflects the overall performance of the economic system as well as various macroeconomic variables of a country. Consequently, forecasting its future behavior (both in the short and long term) has become a major challenge for economic system players and a subject of many scientific research. During the past years, exchange rates were determined by the balance of payments (BoP), which were merely records of a country's receipts and payments in international transactions[1]. Therefore, predicting exchange rates in earlier years was not particularly challenging. Since 1973, with the collapse of the Bretton Woods system of fixed exchange rates and the emergence of the floating exchange rate system in industrial countries, forecasting of exchange rate fluctuations has emerged as a serious challenge. From that time onwards, extensive discussions have taken place regarding exchange rate volatility, its impact on currency, inflation, international trade, and its role in assessing economic security, profitability, risk management, and investment analysis (Ozcelebi, 2018). If we define fluctuation as the risk or uncertainty associated with unpredictable changes in exchange rates over time (Morina et al., 2020), forecasting exchange rates becomes highly challenging due to their nonlinear and volatile nature (Yasir et al., 2019; Huang et al., 2010). Basically, the foreign exchange (Forex) market is a very complicated and volatile market, which is often compared with the Black Box because of its unknown nature and high exchange rate fluctuations (Anastasakis & Mort, 2009). Shocks are the primary source of unpredictable changes, which can impact commodity prices, inflation, interest rates, investment portfolios, savings, and loans (Clarida & Gali, 1994) and ultimately leading to fluctuations in the currency market. Nevertheless, stock volatility and price fluctuations in forex have been proven to be predictable (Sirignano & Cont, 2019).

The stock market is one of the oldest financial markets in the world, and predicting stock prices has been a primary challenge due to the market's nature and behavior. Consequently, extensive efforts and research have been conducted to develop theories, methods or models capable of accurately

[1] Total money coming into a country (inflow) minus total money going out (outflow), consists of the revenue from the export of goods and services minus the costs associated with the import of goods and services (i.e., current account), plus the capital entering the country minus the capital leaving it (i.e., capital account).

[DOI: 10.61882/jme.19.2.249]

predicting stock prices. Some of these theories emphasize the randomness of price fluctuations (Jovanovic & Le Gall, 2001; Fama, 1965; Godfrey et al., 1964; Cootner, 1964; Malkiel, 1999), while others focus on the non-randomness of these changes (Lo, 1997; Lo & MacKinlay, 2011; Lo, 2004). With the collapse of the Bretton Woods system and the gradual establishment of the floating exchange rate system in the Forex market, a similar challenge emerged (this time about the randomness or non-randomness of exchange rate fluctuations). Some theories argue that exchange rate fluctuations are entirely random and that no specific pattern or rule can be defined for them (Rossi, 2013). In contrast, other theories emphasize that past dynamics of the exchange rate can predict its future values (Narayan, 2022). One of the early studies in this field is the research conducted by Meese and Rogoff (1983). Prior to their work, exchange rate prediction methods were concentrated on a limited set of models (Cheung et al., 2019). They employed structural models based on asset and monetary pricing theories for exchange rate prediction, but their model's performance was weak compared to the traditional random walk model for out-of-sample prediction. Nearly two decades of efforts to develop models that could outperform the random walk approach yielded little success. However, Kilian and Taylor (2003) argue that the poor predictive performance of these models is due to the limitations of linear statistical models, while existing studies also confirm the nonlinear behavior of exchange rates. Given the debate between these two perspectives, various methods and models have gradually been developed to confirm or refute this theory, resulting in a broader range of approaches and diversity in the field.

Exchange rate prediction approaches are generally classified into two main categories: fundamental and technical. Within each category, various sub-classifications have been proposed by researchers (Pandey et al., 2018; Sidehabi & Tandungan, 2016; Pandey et al., 2020; Yasir et al., 2019). In general, the term *fundamental* refers to market movements driven by news or factors that can impact a country's economy. Accordingly, the fundamental approach aims to identify the various economic, social, political, and other relevant factors affecting exchange rates, describe their relationships, and formulate predictive models. In contrast, the *technical* approach primarily analyzes market movements by studying charts, price indicators of ongoing market transactions, supply and demand trends, and similar factors (Taylor & Allen, 1992). The fundamental approach can be further divided into two main categories: The first category consists of methods based on economic models, which predict future exchange rate trends using mathematically defined relationships between exchange rate and other fundamental variables, relying

[DOI: 10.61882/jme.19.2.249]

on specific economic theories. The second category includes econometric models, which, while supported by economic theories but are less constrained by theoretical relationships between variables. Instead, they utilize proposed theoretical variables to design the model structure (Rossi, 2013; Orellana & Pino, 2021; Wang et al., 2019; Xu, 2003). Methods related to the technical approach can also be classified into four main categories. The first category includes time series-based models, which primarily analyze the dynamics of the target variable and its associated shocks (Rossi, 2013; Orellana & Pino, 2021; Temür et al., 2019; Engel & Wu, 2023). The second category consists of machine learning-based methods, which are divided into two subgroups: traditional machine learning and deep learning methods (Abiodun et al., 2018; Cao et al., 2005; Oussidi & Elhassouny, 2018; Gu et al., 2023; Temür et al., 2019; Markova, 2019). The third category includes methods that do not fall into the previous two categories (Levy, 1994). Finally, the fourth category comprises hybrid methods, which combines time series analysis with machine learning (Fan et al., 2021; Fathi, 2019; Jamil, 2022; Shui-Ling & Li, 2017; Gu et al., 2023; Temür et al., 2019). Table 1 presents the classification of each approach, along with the methods and models proposed by various researchers.

[DOI: 10.61882/jme.19.2.249]

Table 1
Different exchange rate forecasting approaches, methods, and models

Approach	Category	Method	
fundamental	Economic Models	Uncovered Interest Rate Parity, Purchasing Power Parity, Monetary model with flexible prices, Monetary model with sticky prices, Model with productivity differentials, Portfolio balance model, Taylor rule model, Net foreign asset model, Commodity prices	
	Econometrics Models	Linear Models, Nonlinear Models, Semiparametric Model, Non-parametric Models, Simultaneous Equations, The Panel Data Models	
Technical	Time Series Models	Random Walk (RW), Autoregressive (AR), Moving Average (MA), Autoregressive Moving Average (ARMA), Autoregressive Integrated Moving Average (ARIMA), Seasonal Autoregressive Integrated Moving Average (SARIMA), Autoregressive Fractionally Integrated Moving Average (ARFIMA), Autoregressive conditional heteroskedasticity (ARCH), Generalized Autoregressive Conditional Heteroskedasticity (GARCH), The Error Correction Model (ECM), The Vector Error Correction Model (VECM), Vector Autoregressive (VAR), Threshold Autoregressive Model (TAR), Exponential Smooth Transition Autoregressive Models (ESTAR), Time-Varying Parameter (TVP), Multivariate Time-Varying Parameter (MTVP), Exponential Smoothing (ES), Multivariate Time-Varying Parameter (MTVP)	
	Machine Learning Approaches	Traditional Machine Learning	Naive Bayes, Decision Tree, Random Forest, The Support Vector Machine (SVM), Artificial Neural Network (ANN), Multi-Layer Perceptron (MLP)
		Deep Learning	Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Deep Neural Network (DNN), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Reinforcement Learning (RL)
	Other Approaches	K-Nearest Neighbors algorithm, Chaos Theory, Bayesian Model Averaging (BMA), Linear Discriminant Analysis (LDA)	
	Hybrid	ARIMA+ANN, ARIMA+RNN, ARIMA+ LSTM, ARIMA+SVR, SARIMAX+ LSTM, MF+LSTM, FLB+LSTM, AdaBoost+ LSTM, CNN+RF, CNN+LSTM, CNN+CNN-LSTM, CNN+GRU, GRU+LSTM, RNN+LSTM	

Source: Research findings

Among the various proposed methods, the application of neural networks for exchange rate prediction has been a popular approach over the past two decades. Additionally, advancements in processing capabilities and deep learning-based computational techniques drawing more attention to this field in the last recent years (Ozcan, 2017). A common aspect of most exchange rate prediction studies is the type of data sources used as the primary input for estimation models, which are often based on economic data with time series patterns. Meanwhile, textual data related to currency fluctuations not only generally include quantitative values of fluctuations but also contain

[DOI: 10.61882/jme.19.2.249]

supplementary information that describes the sentiment and intensity of these quantitative data. Such supplementary textual information serves as highly valuable descriptions and insights of data, which, if properly processed, understood and utilized through natural language processing (NLP) techniques and incorporated as additional features in predictive models (particularly deep neural network-based models) can significantly impact the results.

Until now, economic science has focused entirely on designing and developing structures and models based on statistical data and information to explain economic changes. However, there is another influential factor that has been largely overlooked: "Narratives" (Shiller, 2017). The relationship between an economic event such as a crisis or an economic boom and narratives is bidirectional. Sometimes, events give rise to narratives, while at other times, stories and narratives drive economic events. These narratives are expressed in books, historical documents, news media content, social networks, and even large-scale public tweets. Popular and widespread economic narratives have the power to shape people's thinking and decision-making processes by providing them with ideas that influence their perspectives on economic events. As a result, these narratives can significantly impact individuals' economic decisions and actions. People often communicate their values and judgments through storytelling, contributing to the overall sentiment surrounding economic events and trends (Shiller, 2019).

In past years, the lack of necessary tools and technologies for collecting and analyzing stories and narratives has been a key factor in the neglect of their influence on economic events. However, advancements in high-capacity processing technologies, along with the development of data-centric sciences and technologies, have opened a path for narrative analysis. Emerging applications such as "sentiment analysis", "opinion mining", "stance detection" and others are rapidly growing research areas, in the field of data science. In general, sentiment analysis is the automated process of identifying the underlying subjective emotions present in a text. This process involves classifying the opinions expressed in the text based on their polarity, which can be categorized as positive, negative, or neutral (Serrano-Guerrero et al., 2015). Recent advancements in online social networks and the rapid transfer of information have drawn the attention of scientific researchers to sentiment analysis (Naeem et al., 2021). One of the areas where sentiment analysis has gained prominence is in forecasting and estimating the future behavior and trends of financial markets, including the foreign exchange market, based on current market sentiments. Studies have shown that, in the short and medium

[DOI: 10.61882/jme.19.2.249]

term, exchange rates are influenced by market sentiments rather than fundamental factors (Hopper, 1997). Therefore, market sentiment should also be taken into account to ensure accurate predictions (Ozcan, 2017). The integration of sentiment analysis into prediction models has shown a marked improvement in forecasting accuracy, particularly when combined with advanced deep learning architectures. For instance, Jun Gu et al. (2024) developed the FinBERT-LSTM model, which leverages financial news sentiment alongside historical stock data and demonstrates superior performance over baseline models such as DNN and standard LSTM, as evaluated by MAE, MAPE, and accuracy metrics. Similarly, Das et al. (2024) proposed a hybrid EEMD–ensemble CNN model incorporating X (Twitter) sentiment scores, showing enhanced prediction robustness by capturing the public's real-time emotional responses. The use of social media platforms like Twitter further enriches prediction quality, as shown by Sonia et al. (2024), who analysed correlations between sentiment trends and price movements to improve sentiment-driven models. Agrawal et al. (2024) expanded this approach by combining sentiment and technical indicators, resulting in a reinforced predictive framework that adapts across multiple industry sectors. Finally, Shah et al. (2024) highlighted the broader potential of big data sentiment analysis through social and web media, suggesting that emerging machine learning techniques can effectively capture market-relevant emotions, even in the presence of sarcasm or multilingual content. Collectively, these studies underscore the pivotal role of sentiment. Sentiment analysis can be classified into two main approaches: (1) Corpus-based approach, where the emotional affinity of words is determined by learning their probabilistic sentiment scores from large text corpora, and (2) Dictionary-based approach, in which emotionally charged words are automatically extracted using lexical resources such as WordNet[1] and utilized for sentiment classification (Shelke et al., 2012). Sentiment analysis can be performed at word, sentence, document, and aspect level.

Three prominent types of online data sources used for financial analysis include news websites, search engine data, and social media (Ozturk & Ciftci, 2014). Social media platforms contain people's opinions and perspectives on current trends and significant topics, serving as rich sources of unstructured

[1] WordNet is a lexical database for the English language that groups words into sets of synonyms, known as Synsets. It provides concise and general definitions and records the relationships between different synonyms within these sets. The goal of WordNet is to create a combination of a dictionary and a knowledge base that is inherently highly useful and is employed to support automated text analysis and AI applications.

[DOI: 10.61882/jme.19.2.249]

data that can be purposefully used for various applications, including sentiment analysis. A key feature of social media is that individuals from anywhere in the world can freely express their thoughts and beliefs without revealing their real identity or fearing adverse consequences. This anonymity makes these opinions valuable, as they often represent unfiltered and genuine sentiments. However, anonymity in expressing opinions also comes at a cost, as it allows individuals with hidden agendas or malicious intentions to easily spread fake news and misleading information. X (formerly Twitter) is one of the most widely used microblogging social networks, with over 350 million monthly active users, enabling communication through messages of up to 280 characters, known as "tweets" (Duz Tan & Tas, 2021; Antonakaki et al., 2021). Annually, over 500 million Persian tweets are published on X. Due to the short length of tweets, the computational overhead of text processing models is reduced, making X a valuable source of information for researchers. One of the first influential studies in the field of sentiment analysis on X dates back to the work of Go et al., in 2009. Following this research, "sentiment analysis" became one of the three major areas of interest for researchers on X (Antonakaki et al., 2021). The three main advantages of sentiment analysis on X are scalability, real-time analysis, and stability of metrics. "Scalability" enables the analysis of a large number of tweets mentioning a specific topic (Nazir et al., 2019). Additionally, the X data can be labeled continuously to avoid human error-caused inconsistencies (Alrubaian et al., 2018). This paper presents the results of a study focused on predicting the USD/IRR exchange rate using a hybrid model that combines economic variables and sentiment analysis features extracted from Iranian users' tweets in Farsi on the X platform, implemented through a deep neural network. Accordingly, our research question is: Do sentiment analysis features affect the performance of exchange rate prediction models? We hypothesize that sentiment analysis features have an impact on the performance of exchange rate forecasting models.

In the following sections of the paper, Section 2 elaborates on the research methodology in detail, including the methods and tools used throughout the research process. Section 3 presents the results and findings of each research step along with their analysis. Finally, Section 4 discusses the results, summarizes the conclusions, and provides recommendations for future related studies.

[DOI: 10.61882/jme.19.2.249]

2 Methodology and Materials

Overall, the study was conducted simultaneously along two parallel paths, which at the end steps were ultimately integrated. In one path, the process of collecting, preprocessing, and analyzing structured data related to economic variables was carried out using the Cross Industry Standard Process for Data Mining (CRISP-DM) methodology (Schröer et al., 2021; Wirth & Hipp, 2000). In the other path, the collection, preprocessing, and sentiment analysis of dollar-related tweets were performed using the content analysis method (White & Marsh, 2006). Figure 1 illustrates the main stages of the research in these two mentioned paths.

2.1 Tweet Extraction

There are various methods for extracting relevant tweets on a specific topic from Twitter, one of which is the "keyword matching" method (Qi et al., 2020). In this study, to accurately extract Persian tweets related to the Dollar exchange rate, a combination of different keywords such as "dollar", "euro", "gold", "coin", and "currency" was used in the tweet text. However, after conducting various tests, only the two keywords and hashtags "currency" and "dollar" were ultimately chosen for this purpose.

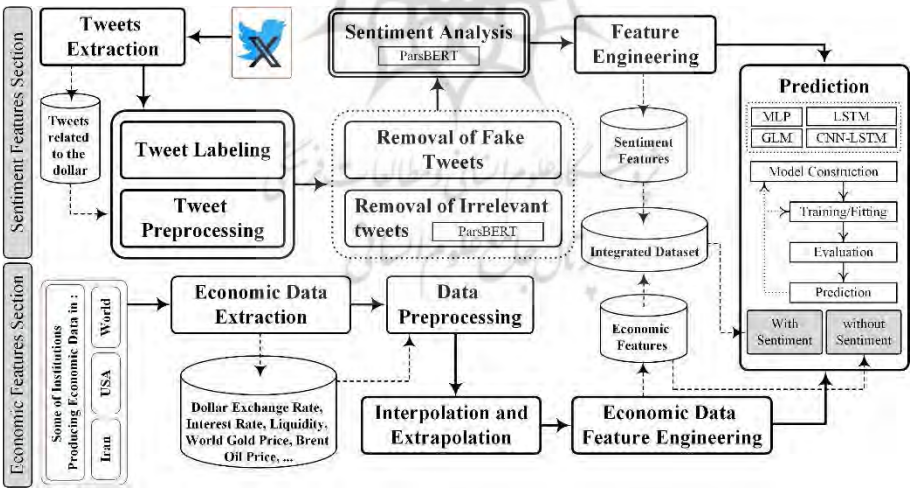

Figure 1. Research methodology and its main steps
Source: Research findings

[DOI: 10.61882/jme.19.2.249]

2.2 Tweet Preprocessing[1]

In general, the content on social networks, which are often unstructured, are neither clean nor standardized and contains many noises. Therefore, to make effective use of them, it is necessary to first enhance their quality (Arolfo et al., 2022; Naeem et al., 2021). Data preprocessing not only has the potential to improve data quality but also helps reduce data volume, which in turn increases the speed of subsequent processing (Ilyas & Rekatsinas, 2022; Lee et al., 2021). Tweet preprocessing was performed in two primary stages: 1) resolving structural errors and data noises and 2) normalizing tweets. To achieve this, a random sample of 3,000 tweets was first extracted to analyze, identify, and categorize different types of errors, noise, and data contamination. Based on this analysis, rules and procedures for error correction were designed, coded, and tested. After developing the preprocessing tool and verifying its functionality, the full dataset of tweets was preprocessed.

2.3 Tweet Labeling

To train the sentiment classification model for dollar-related tweets, a sufficient number of tweets need to be labeled. Generally, sentiment labeling of tweets is performed using two methods: 1) Through a group of experts and 2) Using crowdsourcing techniques. In crowdsourcing techniques, participants are allowed to manually label tweets using online platforms or other annotation tools, usually receiving a small reward in return (Antonakaki et al., 2021). In this study, the second method was employed, and for this purpose, three independent labeling teams were equipped. A coding scheme for tweet labeling was developed, provided and taught to the teams. The labeling decision for each tweet was based on the majority vote among the team members. The labeling process was conducted in two sampling stages, and the samples were distributed to all three teams. The samples were collected prior to any preprocessing of the tweets, ensuring the semantic structure of the tweets text remained intact. Two types of labels were assigned to each tweet: 1) sentiment polarity label, towards the US dollar rate fluctuations (including positive, negative, and neutral) and 2) tweet content type label, (including relevant and irrelevant). A tweet was labeled as

[1] Although in some classifications, actions such as "removing fake tweets" or "removing irrelevant tweets" are also considered as a part of the preprocessing process, in this paper, each has been separately explained to provide a more detailed explanation of their implementation.

[DOI: 10.61882/jme.19.2.249]

"irrelevant" if, despite containing one or both of the words "dollar" and "currency", its symantic content lacked any of the sentiments positive, negative, or even neutral. Training on imbalanced datasets can lead the model to become biased towards the majority class compared to the minority class, which has fewer labeled samples (Susan & Kumar, 2021; Antonakaki et al., 2021). Therefore, to generate a balanced distribution of labels across all classes and allowing each class an equal opportunity of representation during the training process, an algorithm was developed to selectively extract tweets for labeling. The algorithm initially identified time periods when real fluctuations in the exchange rate led to sentiments related to one of the sentiment classes. Given the higher likelihood of publishing tweets with sentiments similar to the real fluctuations during these periods, relevant samples for each class were extracted from these periods. For instance, in Figure 2, the Black, green, and red rectangles represent time periods where the likelihood of extracting neutral, negative, and positive sentiments, respectively, may be higher.

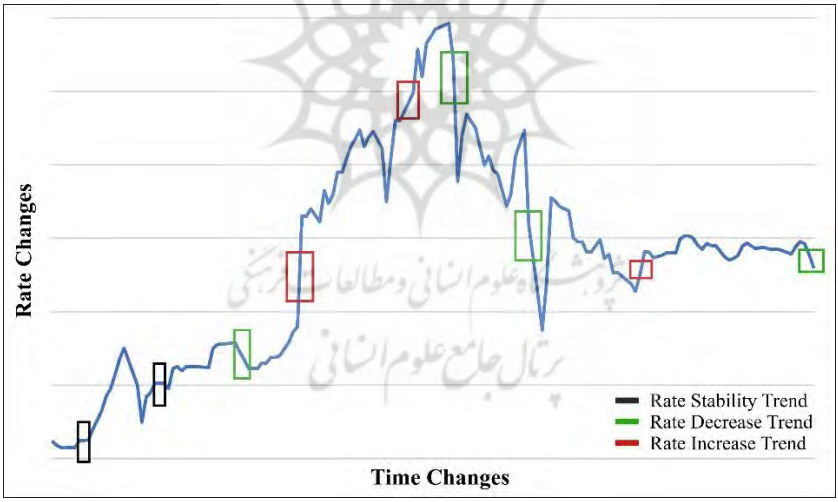

Figure 2. Examples of different types of trends in a representative exchange rate signal
Source: Research findings

2.4 Removal of Invalid Tweets

In general, despite all efforts made in the process of extracting tweets related to a specific topic, the content of many extracted tweets often remains

[DOI: 10.61882/jme.19.2.249]

irrelevant to the subject matter (Patton et al., 2020). Additionally, among the tweets published on specific topics, some are disseminated with purposes other than expressing opinions on the subject, often by fake users or automated agents (Chen et al., 2022). Thus, in a general classification, invalid tweets can be categorized into two main groups: 1) Tweets associated with fake accounts, 2) Tweets unrelated to the topic of exchange rate fluctuations.

2.4.1 Removal of Fake Tweets

The literature reveals multiple methodological approaches for fake tweet identification. Recent studies have employed various deep learning and machine learning techniques for this purpose (Alfonse & Gawich, 2023; Geurgas & Tessler, 2024; Koru & Uluyol, 2024; Mareeswari & Dinesh, 2023; Nikiforos et al., 2020; Sharma et al., 2022). Particularly noteworthy is the work by Monica and Nagarathna (2020), which focused specifically on identification of bots that produce artificial content using deep learning models.

In our study, after a thorough examination of Persian tweets and the distinctive features of fake tweets, we concluded that relying solely on machine learning methods could lead to the removal of a significant number of genuine tweets. Consequently, we adopted an innovative approach that more accurately detects and removes fake tweets (Appendix 1). In total, 228,311 fake tweets were identified and removed from the dataset.

2.4.2 Removal of Irrelevant Tweets

Irrelevant tweets are a subset of extracted tweets that, despite containing the necessary keywords in the extraction phase, do not express opinions related to fluctuations in the US dollar (Garcia-Arteaga et al., 2024). The most common approach for distinguishing between relevant and irrelevant tweets is classification, and among various classification methods, machine learning-based approaches being the most popular among researchers (Antonakaki et al., 2021). In this study, a transfer learning approach was used to remove irrelevant tweets. Figure 3 illustrates the process of eliminating irrelevant tweets.

[DOI: 10.61882/jme.19.2.249]

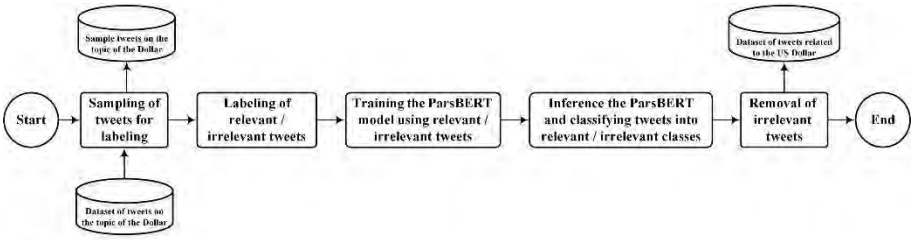

Figure 3. Process of identifying and removing irrelevant tweets
Source: Research findings

In the labeling process, 3,954 tweets were labeled as "relevant" and 7,046 as "irrelevant". For model training, a balanced sample (3,954 relevant and 3,954 irrelevant tweets) was used. The training and validation sample sizes were set at 90% and 10% of the training dataset, respectively. The ParsBERT model was fine-tuned using the labeled tweets for both "relevant" and "irrelevant" classes. After optimizing the model's performance, the F1 scores for the two classes were obtained as 0.79 and 0.77, respectively, with an overall accuracy of 0.78. Subsequently, model inference was performed, classifying 1,631,210 tweets (61%) as irrelevant, which were then removed from the main dataset.

2.5 Dollar Sentiment Analysis

The ParsBERT model was trained using a balanced dataset of 2,322 labeled tweets across three classes: "Positive", "Negative", and "Neutral" (774 labeled tweets for each class). After optimizing model performance by adjusting various hyperparameters, the F1-score values for the three classes were obtained as 0.79, 0.81, and 0.78, respectively, with an overall accuracy of 0.80. The model inference was performed, and the dollar sentiment of each tweet was determined. In this study, the algorithm used for selecting the final class from the probable classes for each tweet was based on choosing the class with the highest probability among the others. For instance, considering the probabilities given in Table 2, the corresponding tweet is classified under the "Negative" class.

[DOI: 10.61882/jme.19.2.249]

Table 2
An example of the probability of a tweet belonging to different sentiment classes

Class Name	Dependency Probability Value
Positive	0.275
Negative	0.673
Neutral	0.052

Source: Research findings

2.6 Dollar Sentiment Feature Engineering

Given that our research objective is to forecast the "daily" dollar exchange rate, the dollar sentiments of individual tweets from a given day must be aggregated into a daily unit. To obtain the "Daily Dollar Sentiment" feature, its numerical value was first calculated from the arithmetic mean of the frequency of classes on a given day. In the computational formula, two parameters, α and β, as well as α±β thresholds, were embedded into its calculation formula to achieve two objectives: first, to compensate for the error caused by excluding the effect of neutral tweets frequency in a given day, and second, to ensure the accurate reflection of the prevailing sentiment of the day's tweets in the selected sentiment class. The optimal values of these two parameters were determined by testing various values and analyzing the correlation between the computed numerical value and the actual daily exchange rate fluctuations (Appendix 1). In the next step, eight features were extracted, with the "numerical value of daily dollar sentiment" chosen as the primary sentiment index for modeling due to several reasons.

Firstly, multicollinearity, a common modeling problem, occurs when two or more explanatory variables exhibit high correlation or linear relationships. In this context, as the "numerical value of daily dollar sentiment" has strong linear relationships and consequently high correlations with other sentiment features, to avoid multicollinearity issues, selecting a single representative feature was necessary. Furthermore, data analysis revealed that during periods of decreasing exchange rates, the number of negative messages did not rise significantly. Instead, there was a decrease in positive messages and an increase in neutral messages. This could be attributed to the Iranian economy's history of rising exchange rates, leading to difficulty in accepting exchange rate decreases. Consequently, simply considering positive, neutral, and negative message counts could result in modeling inaccuracies.

To address this issue, the "numerical value of daily dollar sentiment" was chosen as the primary sentiment index, as it accounts for these unique data conditions. By incorporating this index, the model better reflects the sentiment

[DOI: 10.61882/jme.19.2.249]

of economic actors and accurately represents sentiment analysis in forecasting exchange rates.

2.7 Economic Data Extraction

The data for 24 economic variables in our model were collected from nine different data sources in Iran and the United States over the period from October 2019, to March 2023. These data were produced and published in various formats and standards. Although the earliest observations pertain to the "Dollar Exchange Rate in the Iran Market" and the "Interest Rate in the U.S. Market", dating back to March 22, 1980, the study period was determined based on the first available observation of the "Purchasing Managers Index (PMI)" on October 7, 2019. Additionally, the assumption that extrapolation and interpolation techniques could be applied to repair missing data influenced the selection of this timeframe, resulting in a total of 1,247 days. In Table 3, the "First Published Date" corresponds to the source from which the variable's data were collected for this study.

[DOI: 10.61882/jme.19.2.249]

Table 3
Key Characteristics of the Economic Variables in the Model

Variable	Market	First Publication	publication Frequency	Variable	Market	First Publication	publication Frequency
Exchange Rate	Iran	Jun 1980	Daily	Imports	Iran	Mar 1988	Seasonal
Interest Rate	Iran	Oct 2015	Daily	World Gold Price	World	Jul 2000	Daily
	USA	Jan 1959	Daily, Weekly, Monthly, Annual				
M2 Money Stock	Iran	Mar 2004	Monthly	18K Gold Price	Iran	Jul 2013	Daily
	USA	Jan 1959	Monthly, Seasonal, Annual				
Consumer Price Index (CPI)	Iran	Apr 1982	Seasonal	Gold Coin Price	Iran	Apr 2019	Daily
	USA	Jan 1947	Monthly, quarterly, semi-annually, Annual				
Real Gross Domestic Product (GDP)	Iran	Mar 1991	Seasonal	Brent Oil Price	World	Jul 2007	Minute, Hourly, Daily, Monthly
	USA	Jan 1947	Seasonal, semi-annually, Annual,				
U.S. Dollar Index (DXY)	USA	Jan 1971	Minute, Hourly, Daily, Monthly	TSE Market Index	Iran	Dec 2008	Daily
Purchasing Managers Index (PMI)	Iran	Oct 2019	Monthly	Bitcoin Price	World	Sep 2014	Minute, Hourly, Daily, Monthly
Index of Export-Oriented Stocks	Iran	Mar 2007	Daily	Standard and Poor's 500	USA	Dec 1927	Minute, Hourly, Daily, Monthly
Net Foreign Assets	Iran	Mar 2004	Monthly	US Treasury (5Y)	USA	Jan 1962	Minute, Hourly, Daily, Monthly
Exports	Iran	Mar 1988	Seasonal	Dow Jones Industrial Average	USA	Dec 1991	Minute, Hourly, Daily, Monthly

Source: Research findings

[DOI: 10.61882/jme.19.2.249]

2.8 Economic Data Preprocessing

At this step, base-year normalization for some variables, removal of redundant data, data sorting, structural unification, and overall standardization of economic variable data were all performed.

2.9 Economic Data Interpolation and Extrapolation

Missing data in the time series of certain variables, which occurred due to differences in data publication frequencies (weekly, monthly, or quarterly) or the lack of data on holidays, were addressed using linear interpolation models. Additionally, for some variables, data were missing at the end of the study period, and in these cases, they were extrapolated and estimated using the SARIMA model. At the end of this step, data for all economic variables were made available for every day within the study period (see Appendix 2 for more information).

2.10 Economic Data Feature Engineering

Among the various features used in predictive models (Heaton, 2016) with appropriate performance, 11 types of features (Table 4) were selected and generated for all 23 economic variables. Considering that in the arithmetic equation of some of the selected features, different permutations of the 23 variables were also taken into account, all possible unique permutations of the features were created. At the end of this step, a dataset comprising 20,359 features (including the date, 24 economic variables, and 20,335 new features) for 1,247 days (including all calendar days within the study period) was constructed. The constructed features were scored using five methods: "Information Gain", "Fisher Score", "Variance Threshold", "Dispersion Ratio", and "Random Forest". Feature selection was carried out from the top 100 features across these five methods. Ultimately, the features that scored in a higher number of methods were selected. Accordingly, two features were chosen.

[DOI: 10.61882/jme.19.2.249]

Table 4
Specifications of Candidate Features for Construction, Categorized by Feature Types

Feature Type	Feature Method	Calculation Expression
Differences and Ratios	Difference of two variables	$y = x_1 - x_2$
	Ratio between two simple variables	$y = \frac{x_1}{x_2}$
	Ratio of one variable to the square of another variable	$y = \frac{x_1}{x_2^2}$
	Ratio of two variables multiplications to the square of another variable	$y = \frac{x_1 * x_2}{x_2^2}$
Logarithms and Power Functions	Logarithm of variable	$y = log(x)$
	Second power of variable	$y = x^2$
	Root of variable	$y = \sqrt[2]{x}$
Polynomials	Polynomial of variable	$y = 1 + 5x + 8x^2$
Distance Formula	Euclidean distance between two pairs of variables	$y = \sqrt{(x_1 - x_2)^2 + (x_3 - x_4)^2}$
Rational Differences and Polynomials	Four variables difference ratio	$y = \frac{x_1 - x_2}{x_3 - x_4}$
	Variable polynomial ratio	$y = \frac{1}{5x + 8x^2}$

Source: Research findings

2.11 Dollar Exchange Rate Forecasting

During the data preparation process, two datasets with identical structures were created. The first dataset included all economic variable data along with the constructed features (a total of 26 data items). In the second dataset, in addition to the aforementioned data, the numerical value of "daily dollar sentiment" was also incorporated, bringing the total to 27 data items. To enhance the generalizability of the results, four type of prediction models were employed: a linear method (GLM), a traditional machine learning approach (MLP), and deep learning techniques (LSTM and CNN-LSTM). Assuming the two input datasets (with and without dollar sentiment), a total of 8 models were developed. The main steps in the forecasting process were as follows:

- **Model Development:** In this stage, the key hyperparameters of the model were determined and adjusted. Different hyperparameters were set for each of the model types. For instance, in deep learning-based models, some of these hyperparameters included the number of hidden layers, the number of units per layer, the activation function, the optimization algorithm, the learning rate, the number of epochs, and the batch size.
- **Model Training/Fitting:** The constructed model was fitted using the training data. This process ensured that the model parameters were optimally adjusted based on the patterns and relationships learned from

[DOI: 10.61882/jme.19.2.249]

the training dataset. Examples of such parameters include the connection weights in deep learning-based models and the variable coefficients in Flinear models.

- **Prediction:** At this step, the trained and optimized model with the validation data, was forecasted the values of the target variable for the test dataset.
- **Model Evaluation**: At this step, the deviation of the predicted values from the test data was measured. Since this study aims to predict the numerical value of the target function, MAE and MSE were used as evaluation metrics. If the improvement in model performance, improvement was less than the specified threshold, the process was repeated. Otherwise, the optimization procedure was finalized.

3 Results

3.1 Dataset

In this study, two main datasets were prepared and integrated to yield a dataset containing input data for the models.

3.1.1 Tweet Dataset

Based on the keywords, 2,906,045 tweets were extracted. For each tweet, 9 data items were collected (Table 5).

Table 5
Data items collected for each tweet

Data Name	**Tweet Time**	**Tweet Text**	**Account Name**	**User Name**	**Number of Followers**	**Number of Retweets**	**Number of Likes**	**Hashtags**	**Tweet Address**
Data Type	Time Stamp	String	String	String	Number	Number	Number	String	String

Source: Research findings

Statistical analyses of tweets indicate that the temporal pattern of tweet publication corresponds to the Iran's local time (UTC +3:30) (Figure 4). In other words, the highest number of tweets were published during peak working hours and also late evening hours. Furthermore, the volume of tweets published during the weekend days in Iran (i.e., Thursday and Friday) significantly decreased compared to other days of the week (Figure 5).

[DOI: 10.61882/jme.19.2.249]

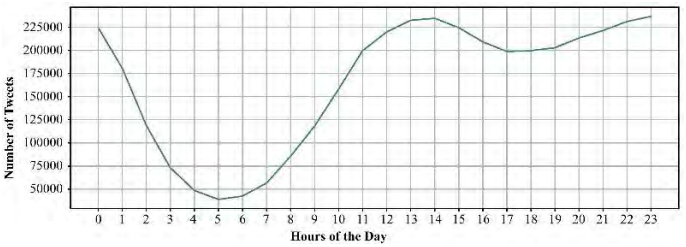

Figure 4. Distribution of the number of tweets based on tweet post time
Source: Research findings

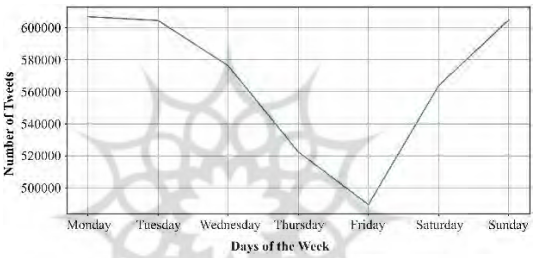

Figure 5. Distribution of the Number of Tweets Across Days of the Week
Source: Research findings

3.1.2 Economic Variables Dataset

We gathered raw time-series data for 24 economic variables from a variety of data-providing sources (Appendix 3) using various methods, including manual data entry from printed reports. After the data collection process, we created a CSV file with 25 columns, comprising 24 columns for the time-series data of the economic variables and one column for the corresponding dates, spanning a total of 15,690 days. Notably, many cells in the file were missing data.

3.2 Tweet Labeling

Using the described algorithm for sample preparation for labeling, we initially retrieved a sample of 3,000 tweets and labeled them. The results indicated that the developed algorithm for targeted tweet extraction needed optimization.

[DOI: 10.61882/jme.19.2.249]

After optimizing, the second sample consisting of 8,000 tweets was extracted and labeled. The results of the labeling process are presented in Table 6.

Table 6
Results of Preparation and Labeling of Tweet Samples

	Label Type	Before Labeling	After Labeling	
		Sample Size	**Number of labeled**	**Percentage**
First Sample	Negative	1000	92	30.7
	Neutral	1000	545	18.17
	Positive	1000	754	25.13
	Unrelated	-	1609	53.63
Second Sample	Negative	3500	682	8.53
	Neutral	1500	382	4.77
	Positive	3000	1499	18.74
	Unrelated	-	5437	67.96

Source: Research findings

3.3 Data Preprocessing

3.3.1 Tweet Preprocessing

Tweet cleaning and normalization process was performed using a custom developed tool, using the ready-made "Hazm[1]" library. Table 7 presents the types of identified issues from the sample review phase, the corrective actions taken, and the number of cleaned tweets in the dataset.

[1] HAZM is a Python library for Persian language processing, which performs tasks such as normalization, sentence and word extraction, stemming, morphological and syntactic analysis, and the identification of syntactic dependencies in Persian texts. HAZM has been developed based on the NLTK (Natural Language Toolkit) library and has been adapted for the Persian language.

[DOI: 10.61882/jme.19.2.249]

Table 7
Tweet preprocessing results

Type of error	Description of error	Measure taken to fix the error	No. of tweets
Irrelevant characters	Characters such as phone numbers, email addresses, double spaces, currency acronyms (symbols), internet addresses, line breaks, and so on.	Removing irrelevant characters	556,529
Excessive tags and unnecessary characters	Tags like CSS, HTML, JS, abbreviations, hashtags, and other similar items	Removing unnecessary tags and characters	588741
Texts with a combination of different languages, where Persian is not the dominant language	Texts with over 20% of their content contain non-Persian characters	Removing the text	748
The presence of ASCII characters	The presence of unreadable ASCII characters in Persian texts	Removal of ASCII characters	1,588,829
Arabic characters with Persian equivalents	The presence of similar Arabic characters in Persian texts like ی and ک in Persian. At the time of importing data, their similar Arabic words, i.e., ی and ک, are inserted in the text by mistake through the Arabic keyboard.	Replacing Arabic characters with Persian equivalents	325,164
Unnecessary or non-standard unicode characters	Characters such as |, \n, --, (…, …), etc., as well as useless Unicode characters that lack semantic value, such as \u2066 (left-to-right)	Deleting some and replacing others with appropriate characters such as "(" or ")"	457,264
Non-normalized Persian text	Some examples of non-normalized text include extra spaces, English/Arabic numbers, unknown punctuation	Normalizing the text using the "Hazm" library	1,617,527

Source: Research findings

3.3.2 Economic Variables Preprocessing

Table 8 summarizes the key actions taken for cleansing, standardizing, and harmonizing the structure of economic variable data, which were collected from various sources with differing formats and content structures. Additionally, the table specifies the number of variables for which each action was applied.

[DOI: 10.61882/jme.19.2.249]

Table 8
Economic variables data preprocessing results

Type of Issue	No. of variables
Removal of redundant records from the beginning of some data	2
Adding a column for the equivalent base date of the collected data (in the Jalali or Gregorian calendar) and completing it based on the reference date	24
Verifying consistency between the variable's date and the exchange rate date (target variable)	10
Converting the effective interest rate (EIR) to the nominal interest rate (NIR) due to differences in data production standards between Iran and the U.S.	1
Unification of the base year for some variables (2011)	3

Source: Research findings

3.4 Irrelevant Tweet Removal

Based on the methodology described in Appendix 4 and the rules of thumb formulated for detecting and removing fake tweets, a Python program was developed and executed. Table 9 presents the results of applying these rules along with their execution order.

Table 9
Results of the execution of the fake tweet removal program

Execution Order	Rule Title	No. of tweets removed
1	Elimination of tweets associated with user accounts containing the term "bot"	2813
2	Removal of duplicate tweets of a user account with more than 2 repetitions	37183
3	Removal of the second instance of tweets from user accounts with only two repetitions	15043
4	Removal of tweets linked to user accounts with non-conventional publication patterns	173272
Total fake tweets removed		228311

Source: Research findings

The model inferred the classification of related and unrelated tweets, and subsequently, unrelated tweets were removed. Ultimately, the changes resulting from this process and the fake tweet removal procedure led to a reduction in the tweet dataset size, decreasing from 2,906,045 tweets to 1,046,524 tweets (Table 10).

[DOI: 10.61882/jme.19.2.249]

Table 10
Changes in the tweet dataset during the irrelevant tweet removal process

Step Title	Dataset Type	No. of tweets	Description
Data collection	All tweets	2906045	Total tweets collected
Removal of invalid tweets	Fake tweets	228311	Tweets removed from the dataset
	Irrelevant tweets	1631210	Tweets removed from the dataset
Number of dataset tweets		1046524	No. of tweets in the dataset for sentiment analysis

Source: Research findings

3.5 Results of Tweet Sentiment Analysis

After performing the sentiment classification model inference for tweets into three classes: "positive", "negative", and "neutral", the results shown in Table 11 were obtained. As a result of the actions in this step, a new basic feature called "tweet dollar sentiment", was added to the tweet dataset. Although this feature's value was obtained for each tweet, and in the final prediction model of this study, all variables and features were assumed on a daily basis, this feature served as the basis for constructing other sentiment-related features considered in the final prediction model.

Table 11
Tweets sentiment distribution in the three classes

Class Name	Class Numerical label	No. of tweets in Class	Percentage
Positive Tweet Sentiment	+1	421979	40
Negative Tweet Sentiment	-1	129349	12
Neutral Tweet Sentiment	0	495196	48
Total Number of tweets		1046524	100

Source: Research findings

3.6 Results of Economic Data Interpolation and Extrapolation

During the data enrichment process, time series data for 21 economic variables were interpolated, and 11 variables were extrapolated. For 10 variables, both processes were performed. The execution of these two processes addressed the missing data in the time series of economic variables. As a result, by the end of this step, the values for all variables were available for every day within the study period (1247 days). The optimal values obtained for the SARIMA model parameters, as well as the Akaike Information Criterion (AIC) value obtained during the model parameter optimization process are presented for each extrapolated variable in Table 12 (further information is in Appendix 2).

[DOI: 10.61882/jme.19.2.249]

Table 12
Parameters and AIC of the selected models for some of economic variables extrapolation

No.	Variable	parameter values ARIMA(*p,i,q*)	parameter values SARIMA(*p,i,q,s*)	AIC value
1	Interest Rate in the US market	(1,0,0)	(2,1,2,14)	-3895.91
2	M2 Money Stock in the Iran market	(2,0,0)	(0,0,0,14)	6
3	M2 Money Stock in the US market	(2,0,0)	(0,0,0,14)	6
4	Consumer Price Index in the Iran market	(1,1,0)	(0,0,0,14)	-3565.33
5	Consumer Price Index in the US market	(1,1,0)	(0,0,0,14)	-8829.06
6	Real Gross Domestic Product in the Iran market	(1,0,0)	(1,1,1,14)	8
7	Real Gross Domestic Product in the US market	(0,0,0)	(2,0,0,14)	6
8	Purchasing Managers Index	(1,1,0)	(0,0,0,14)	-2854.77
9	Net Foreign Assets	(2,0,0)	(0,0,0,14)	6
10	Exports	(1,0,0)	(1,1,0,14)	6
11	Imports	(1,0,0)	(1,1,0,14)	6

Source: Research findings

3.7 Results of Models Construction and Prediction

3.7.1 Models Construction

LSTM

In this study, we utilized a carefully designed LSTM model architecture comprising input and output layers, along with three hidden layers. The number of nodes in each hidden layer was determined using a rule of thumb, which suggests setting the number of nodes to approximately two-thirds of the total number of input and output nodes. Consequently, for both models (the model incorporating dollar sentiment feature and the model excluding dollar sentiment feature), we employed 18 nodes in each hidden layer. This approach allowed us to maintain a balanced complexity within the network, ensuring that it had sufficient capacity to capture intricate patterns in the data while avoiding overfitting. The resulting architecture demonstrated its effectiveness in achieving accurate and robust time series forecasting performance (Figure 6).

[DOI: 10.61882/jme.19.2.249]

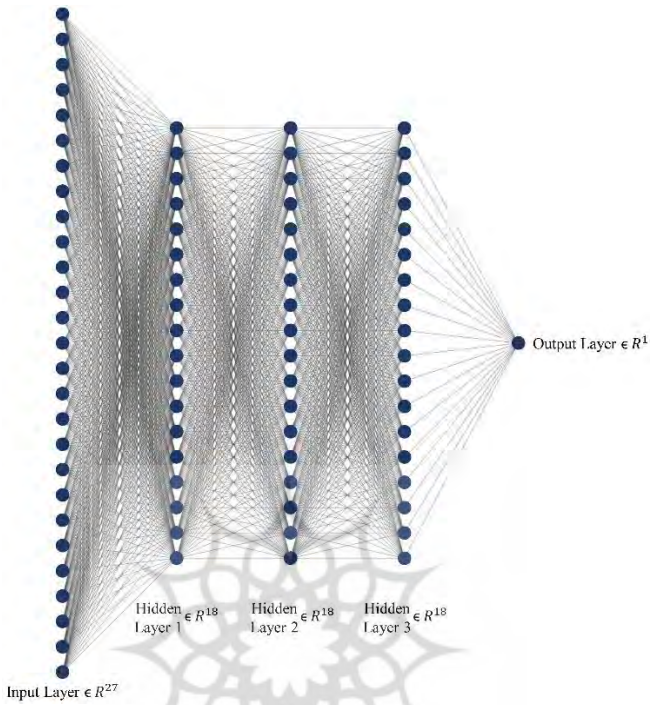

Figure 6. Architecture of the LSTM Model for Data Simulation and Prediction.
Source: Research findings

Additionally, 20% of the dataset was designated for model training purposes, while an additional 20% was reserved for validation. This strategic division allowed us to evaluate the model's performance on unseen data during the training process, enabling us to fine-tune the model's hyperparameters and optimize its predictive capabilities. By maintaining a separate validation set, we ensured that our model's performance assessment remained unbiased and that any potential overfitting issues could be promptly identified and addressed.

In order to identify the optimal LSTM model configuration for our task, we performed a comprehensive hyperparameter tuning process. We focused on five key hyperparameters: activation function, lookback, learning rate, optimization algorithm, and batch size. Various values were considered for each hyperparameter, as presented in the table 13.

[DOI: 10.61882/jme.19.2.249]

Table 13
Proposed Hyperparameter Ranges for Optimization of the LSTM Model.

Hyper Parameters	Considered Values
Activation Function	Tanh, Sigmoid, Relu
Lookback	5, 10, 15, 20, 25, 30
Optimization Algorithm	Adam, Nadam Adamax, Adadelta, SGD, RMSprop
Learning Rate	0.01, 0.001, 0.0001, 0.00001
Batch size	4, 8, 16, 32, 64, 128

Source: Research findings

After conducting an extensive hyperparameter optimization process, we identified the optimal LSTM model that yielded the lowest Mean Square Error (MSE) on our validation dataset. The MSE was chosen as the loss function due to its appropriateness for regression tasks, as it assesses the average squared difference between the predicted and actual values. The table below presents the selected hyperparameters that constitute the best-performing LSTM model architecture:

Table 14
LSTM Configuration of Best Models (with and without Sentiment Features)

Hyper Parameters	Model With Sentiment Feature	Model Without Sentiment Feature
Train/Test split	80/20	80/20
Look Back	5	5
Number of Units	18	18
Number of Hidden Layer	3	3
Act Fun	Tanh	Tanh
Dropout	0.2	0.2
Optimizer	Nadam	Adadelta
Loss function	MSE	MSE
Learning Rate	0.001	0.001
Epoch	100	100
Validation Split	0.2	0.2
Batch size	32	32

Source: Research findings

This fine-tuned LSTM model was ultimately selected for our experiments, ensuring reliable predictions and generalization capabilities on our time series forecasting task.

CNN-LSTM

In this study, we developed a hybrid CNN-LSTM model for data simulation and prediction. To enable compatibility with the CNN component of the

[DOI: 10.61882/jme.19.2.249]

model, an additional dimension was added to the data, resulting in a 4-dimensional input. This preprocessed data was then fed into the convolutional layer, where 27 distinct convolution filters were applied to extract relevant features.

Subsequently, the output of the convolutional layer was reshaped by a flattening layer, reducing its dimensionality from 4 to 3. This transformation prepared the data for subsequent processing by the LSTM component of the model. Given the nature of the data and the objective of leveraging all available information, a pooling layer was intentionally excluded from the CNN architecture.

The flattened output was then passed to an LSTM model consisting of three hidden layers, which captured sequential dependencies within the data. Finally, the output of the LSTM model provided the predicted results. This hybrid CNN-LSTM approach demonstrated its effectiveness in capturing both spatial and temporal patterns within the data, ultimately leading to accurate and robust predictions (Figure 7).

In the CNN-LSTM model, we performed hyperparameter optimization to identify the optimal configuration. We considered various values for five key hyperparameters: activation function, lookback, learning rate, optimization algorithm, and batch size. These hyperparameters and their respective values were the same as those suggested for the standalone LSTM model, as indicated in Table 14. Following the preparation of the data, it was fed into the model for training and validation. The performance of the various model configurations was then assessed using the Mean Square Error (MSE) function. Based on the MSE evaluations, the optimal hyperparameters were identified, which in turn led to the selection of the best-performing model.

Table 15 summarizes the optimized hyperparameters for the best-performing models, including both the model with sentiment feature and the model without sentiment feature. The values listed in the table highlight the optimal configuration for each model variant, providing a direct comparison of the key hyperparameters that contributed to their superior performance.

[DOI: 10.61882/jme.19.2.249]

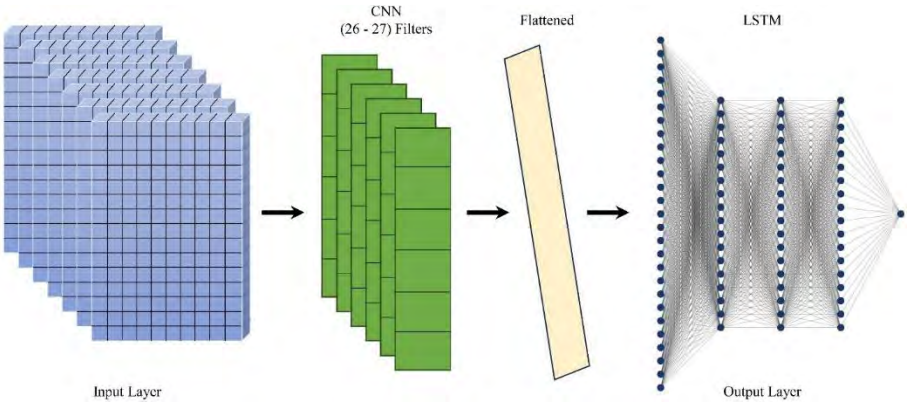

Figure 7. Architecture of the Hybrid CNN-LSTM Model for Data Simulation and Prediction.
Source: Research findings

Table 15
CNN-LSTM Configuration of Best Models (with and without Sentiment Features)

Hyper Parameters	**Model With Sentiment Feature**	**Model Without Sentiment Feature**
Train/Test split	80/20	80/20
Look Back	25	15
Number of Units	18	18
Number of Convolution Layer (CNN)	1	1
Number of Hidden Layer (LSTM)	3	3
Act Fun	Tanh	Tanh
Dropout	0.2	0.2
Optimizer	RMSprop	Nadam
Loss function	MSE	MSE
Learning Rate	0.001	0.01
Epoch	100	100
Validation Split	0.2	0.2
Batch size	16	32

Source: Research findings

MLP

In order to establish the model architecture, it was necessary to determine the size of hidden layer. We employed a rule of thumb to obtain the maximum

[DOI: 10.61882/jme.19.2.249]

size of hidden layer, which involved calculating two-thirds of the combined total of input and output layers, and then adding one.

To identify the optimal hyperparameters for our model, we utilized the dual annealing global optimization algorithm, which allowed us to minimize the loss function within a four-dimensional space. We explored an extensive search range for three key hyperparameters: the hidden layer size (ranging from 1 to 22), random seed (ranging from 1 to 30), and learning rate (ranging from 0.00001 to 1), as presented in Table 16. This comprehensive optimization approach ensured that our model achieved the best possible performance in terms of prediction accuracy and generalization capabilities.

Table 16
Proposed Hyperparameter Ranges for Optimization of the MLP Model.

Hyper Parameters	**Considered Values**	**Initial Values**
Hidden layer size	{1, 2, . . ., 22}	1
Random seed	{1, 2, . . ., 30}	1
Learning rate	[0.00001:1]	0.00001

Source: Research findings

Figure 8 illustrates the comprehensive architecture of the model, incorporating sentiment analysis features and a 22-node hidden layer.

[DOI: 10.61882/jme.19.2.249]

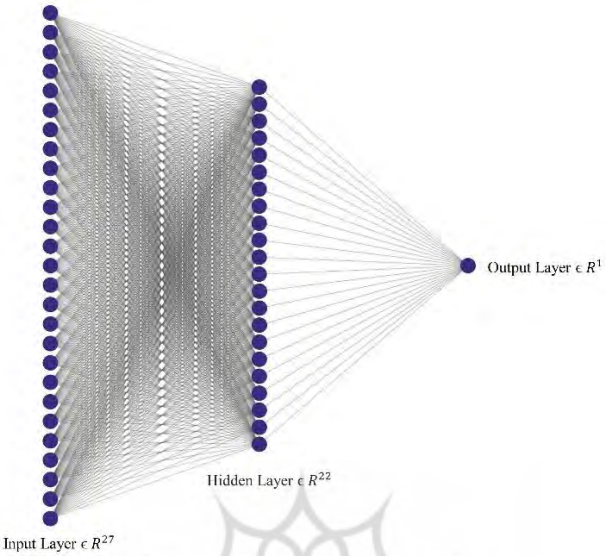

Figure 8. Architecture of the MLP Model for Data Simulation and Prediction.
Source: Research findings

Table 17 highlights that the most effective model configuration includes 1 node in the hidden layer when incorporating dollar sentiment features, and 4 nodes in the hidden layer when excluding sentiment features. Both models achieved the minimum loss value with a random seed of 1. Additionally, a learning rate of approximately 0.0793 was found to be optimal for both models, ensuring efficient convergence during the training process.

[DOI: 10.61882/jme.19.2.249]

Table 17
MLP Configuration of Best Models (with and without Sentiment Features)

Hyper Parameters	Model With Sentiment Feature	Model Without Sentiment Feature
Train/Test split	80/20	80/20
Hidden layer size	1	4
Global optimization algorithm	Dual annealing	Dual annealing
Loss function	MSE	MSE
Learning Rate	0.07932	0.07935
Random seed	1	1
Validation Split	0.2	0.2

Source: Research findings

GLM

A key objective of this research is to investigate whether linear models, such as the Generalized Linear Model (GLM), can accurately capture the impact of sentiment characteristics on exchange rate forecasting performance. To address this question, we employed a GLM model as a representative of linear models for simulating and predicting exchange rates. GLMs extend the simple linear model by allowing the target variable to depend on multiple explanatory variables. A distinguishing feature of GLMs is their inclusion of a random component, which specifies the probability distribution of the target variable; in our study, we selected the Gaussian distribution.

Another essential element of GLMs is the link function, which describes the relationship between the expected value of the target variable and the linear predictor (a linear combination of the explanatory variables). Since our target variable consists of continuous data and was assumed to follow a Gaussian distribution, the identity link function was selected for our analysis. This choice of link function ensures that the expected value of the target variable maintains the same scale as the linear predictor, making it suitable for modeling continuous data with a Gaussian distribution.

3.7.2 Results and discussion

The results reported in table 18 reveal that incorporating sentiment analysis features led to improved predictive performance across all examined models, including LSTM, CNN-LSTM, MLP, and GLM. This finding suggests that integrating sentiment analysis of market participants can enhance exchange rate prediction accuracy, even when accounting for the various economic factors that influence exchange rates. Consequently, sentiment analysis may serve as a valuable tool for refining exchange rate forecasting models and better understanding the complex dynamics of foreign exchange markets.

[DOI: 10.61882/jme.19.2.249]

Table 18
Comparison of Models Evaluation Results (with and without Sentiment Feature)

Evaluation / Models	With sentiment feature		Without sentiment feature	
	MAE	MSE	MAE	MSE
LSTM	18695.89	6.07E+08	22620.95	8.84E+08
CNN-LSTM	43772.96	2.83E+09	44153.71	3.39E+09
MLP	8094.92	1.17E+08	10219.5	1.68E+08
GLM	6630.62	7.25E+07	6724.65	7.39E+07

Source: Research findings

In the context of narrative economics, the observed improvement in exchange rate prediction performance when incorporating sentiment analysis feature can be interpreted as a demonstration of the significant role that narratives and subjective beliefs play in influencing economic outcomes.

Narrative economics posits that economic fluctuations are driven not only by objective factors such as changes in fundamentals, policies, or institutions, but also by the spread and evolution of popular narratives that shape people's expectations, preferences, and actions. In the case of exchange rate forecasting, market participants' sentiment and expectations can be viewed as manifestations of such narratives, reflecting their beliefs about future economic developments and market conditions.

The finding that sentiment analysis can improve exchange rate prediction models suggests that these subjective factors have a tangible impact on the behavior of exchange rates, beyond the traditional economic variables typically considered. This lends empirical support to the narrative economics perspective, highlighting the importance of understanding and accounting for the role of narratives and beliefs in shaping economic phenomena.

In the context of market efficiency and the randomness of exchange rates, as proposed by Meese and Rogoff (1983), our results offer interesting insights. Meese and Rogoff's (1983) study found that fundamental economic factors could not reliably predict exchange rate movements and concluded that exchange rates exhibit random walk behavior.

Our findings, however, suggest that incorporating sentiment analysis feature into predictive models can enhance the accuracy of exchange rate forecasting. This implies that sentiment analysis may capture information that is not reflected in traditional economic factors, potentially providing a competitive edge in predicting exchange rates.

Given these results, it could be argued that exchange rates are not entirely random or efficient in the traditional sense. Instead, the sentiment and

[DOI: 10.61882/jme.19.2.249]

expectations of market participants, which are often based on subjective narratives, may influence exchange rate movements in a somewhat predictable manner.

This does not necessarily contradict the Meese and Rogoff (1983) puzzle entirely but rather indicates that incorporating additional information, such as sentiment data, could improve our understanding and forecasting of exchange rates. In this sense, the efficiency of the foreign exchange market may depend on the availability and incorporation of relevant sentiment information into market participants' decision-making processes.

Our results indicate that the linear model (GLM) outperforms the more complex LSTM and CNN-LSTM models in predicting the rial-dollar exchange rate. This suggests that the rial-dollar exchange rate can be more accurately estimated using a linear model, which implies that the relationship between the predictors and the exchange rate may be linear in nature.

On the other hand, the LSTM and CNN-LSTM models did not perform as well, potentially due to their inherent long-term memory components. In the context of the Iranian exchange rate, which is frequently subjected to new shocks and persistent increases, past values may hold limited predictive power. Consequently, the deep learning memory models' ability to retain historical information may not significantly contribute to enhancing forecast accuracy in this particular case. These findings highlight the importance of considering the specific characteristics and dynamics of the target variable when selecting an appropriate forecasting model.

4 Conclusion

In conclusion, this study investigated the potential of various machine learning and linear models for predicting the rial-dollar exchange rate, with a particular focus on the impact of incorporating sentiment analysis feature. Our findings suggest that the GLM, as a linear model, provided the most accurate forecasts, outperforming more complex deep learning models such as LSTM and CNN-LSTM. This indicates that the relationship between the predictors and the rial-dollar exchange rate may be linear in nature.

In contrast, the LSTM and CNN-LSTM models did not yield superior performance compared to the GLM, potentially due to the Iranian exchange rate's unique characteristics and susceptibility to new shocks. The long-term memory component of these deep learning models may not offer significant benefits in this context, as past values may hold limited predictive power for the exchange rate's future behavior.

[DOI: 10.61882/jme.19.2.249]

Our research provides important implications for narrative economics and the broader understanding of exchange rate dynamics. By examining the role of sentiment analysis feature in enhancing the predictive performance of various forecasting models, our study highlights the significance of narratives and subjective beliefs in influencing economic outcomes.

Narrative economics posits that economic fluctuations are driven not only by objective factors such as changes in fundamentals, policies, or institutions, but also by the spread and evolution of popular narratives that shape people's expectations, preferences, and actions. In the context of exchange rate forecasting, sentiment analysis can be viewed as a means to capture these narratives, reflecting market participants' beliefs about future economic developments and market conditions.

Our results demonstrate that incorporating sentiment analysis leads to improved exchange rate predictions, even when accounting for traditional economic variables. This finding lends empirical support to the narrative economics perspective, underscoring the importance of understanding and considering the role of narratives and subjective factors in shaping exchange rate dynamics.

The implications of our research extend beyond the realm of exchange rate forecasting, as they emphasize the need to account for the influence of narratives and subjective beliefs in economic models more broadly. By recognizing the significant impact of these factors, researchers and policymakers can develop more comprehensive and accurate models, ultimately leading to better-informed decision-making processes and more effective policies.

Despite the valuable contributions of our research, several limitations should be acknowledged. Firstly, the restricted access to Twitter's infrastructure services within Iran posed challenges in collecting bulk tweet data. Secondly, the diverse sources of economic data presented difficulties in standardizing and harmonizing the information gathered from multiple sources. While some data sources covered a broader time range than the scope of our study, it was necessary to limit the analysis to periods where data was available for all relevant variables. Consequently, potentially valuable data from certain sources had to be excluded, which may have influenced the robustness of our findings. Lastly, due to constraints on research time and resources, it was not feasible to expand the scope of our study to include a larger volume of labeled data. This limitation may have prevented our models from achieving even higher levels of accuracy and performance.

[DOI: 10.61882/jme.19.2.249]

Future research directions can be categorized into two main areas: enhancing data preprocessing techniques and expanding sentiment analysis.

In terms of data preprocessing, exploring advanced methods for identifying and removing irrelevant or invalid tweet data can significantly improve data quality. This could involve employing more sophisticated machine learning approaches or refining existing algorithms. Additionally, conducting a comprehensive feature engineering process, although resource-intensive, could lead to a more robust model with improved performance.

Expanding sentiment analysis is another promising avenue for future research. This includes performing multilevel sentiment analysis, such as at the document or aspect level, and utilizing alternative classification techniques. Moreover, extending sentiment analysis to other sources like news websites, digital marketplaces, and additional social networks can enhance sentiment extraction quality. By addressing these areas, future studies can contribute to the development of more accurate and reliable exchange rate forecasting models that account for a broader range of factors and data sources.

Furthermore, future research could further explore the potential of hybrid models or alternative deep learning architectures tailored to the unique features of exchange rates, as well as investigate the impact of other subjective factors on exchange rate forecasting.

References

Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A., & Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. *Heliyon*, *4*(11).

Agrawal, S., Kumar, N., Rathee, G., Kerrache, C. A., Calafate, C. T., & Bilal, M. (2024). Improving stock market prediction accuracy using sentiment and technical analysis. *Electronic Commerce Research*, 1-24.

Alfonse, M. A. R. C. O., & Gawich, M. (2023). Emotions-Based Disaster Tweets Classification: Real or Fake. *WSEAS Transactions on Information Science and Applications*, *20*, 313-321.

Alrubaian, M., Al-Qurishi, M., Alamri, A., Al-Rakhami, M., Hassan, M. M., & Fortino, G. (2018). Credibility in online social networks: A survey. *IEEE Access*, *7*, 2828-2855.

Anastasakis, L., & Mort, N. (2009). Exchange rate forecasting using a combined parametric and nonparametric self-organising modelling approach. *Expert Systems with Applications*, 36, 12001–12011.

[DOI: 10.61882/jme.19.2.249]

Antonakaki, D., Fragopoulou, P., & Ioannidis, S. (2021). A survey of Twitter research: Data model, graph structure, sentiment analysis and attacks. *Expert systems with applications*, *164*, 114006.

Arolfo, F., Rodriguez, K. C., & Vaisman, A. (2022). Analyzing the quality of Twitter data streams. *Information Systems Frontiers*, *24*(1), 349-369.

Cao, D. Z., Pang, S. L., & Bai, Y. H. (2005). Forecasting exchange rate using support vector machines. *In 2005 international conference on machine learning and cyber*netics (Vol. 6, pp. 3448-3452). IEEE.

Chen, K., Duan, Z., & Yang, S. (2022). Twitter as research data: Tools, costs, skill sets, and lessons learned. *Politics and the Life Sciences*, *41*(1), 114-130.

Cheung, Y. W., Chinn, M. D., Pascual, A. G., & Zhang, Y. (2019). Exchange rate prediction redux: New models, new data, new currencies. *Journal of International Money and Finance*, *95*, 332-362.

Clarida, R., & Gali, J. (1994). Sources of real exchange-rate fluctuations: How important are nominal shocks? In *Carnegie-Rochester conference series on public policy* (Vol. 41, pp. 1-56). North-Holland.

Cootner, P. H. (1964). The random character of stock market prices. MIT Press.

Das, N., Sadhukhan, B., Bhakta, S. S., & Chakrabarti, S. (2024). Integrating EEMD and ensemble CNN with X (Twitter) sentiment for enhanced stock price predictions. *Social Network Analysis and Mining*, *14*(1), 29

Duz Tan, S., & Tas, O. (2021). Social media sentiment in international stock returns and trading activity. *Journal of Behavioral Finance*, *22*(2), 221-234.

Engel, C., & Wu, S. P. Y. (2023). Forecasting the US Dollar in the 21st Century. *Journal of International Economics*, *141*, 103715.

Fama, E. F. (1965). The behavior of stock-market prices. *The journal of Business*, *38*(1), 34-105.

Fan, D., Sun, H., Yao, J., Zhang, K., Yan, X., & Sun, Z. (2021). Well production forecasting based on ARIMA-LSTM model considering manual operations. *Energy*, *220*, 119708.

Farahani, M., Gharachorloo, M., Farahani, M., & Manthouri, M. (2021). Parsbert: Transformer-based model for persian language understanding. *Neural Processing Letters*, *53*, 3831-3847.

Fathi, O. (2019). Time series forecasting using a hybrid ARIMA and LSTM model. *Velvet Consulting*, 1-7.

Garcia-Arteaga, J., Zambrano-Zambrano, J., Parraga-Alava, J., & Rodas-Silva, J. (2024). An effective approach for identifying keywords as high-quality filters to get emergency-implicated Twitter Spanish data. *Computer Speech & Language*, *84*, 101579.

Geurgas, R., & Tessler, L. R. (2024). Automatic detection of fake tweets about the COVID-19 Vaccine in Portuguese. *Social Network Analysis and Mining*, *14*(1), 55.

Go, A., Bhayani, R., & Huang, L. (2009). Twitter sentiment classification using distant supervision. *CS224N project report, Stanford*, *1*(12), 2009.

[DOI: 10.61882/jme.19.2.249]

Godfrey, M. D., Granger, C. W., & Morgenstern, O. (1964). The Random Walk Hypothesis Of Stock Market Behavior a. *Kyklos*, *17*(1), 1-30.

Gu, R., Du, Y., Guo, S., Zhang, Z., & Wu, K. (2023). RMB exchange rate forecasting algorithm and empirical analysis. In *3rd International Conference on Management Science and Software Engineering (ICMSSE 2023)* (pp. 96-104). Atlantis Press.

Heaton, J. (2016). An empirical analysis of feature engineering for predictive modeling. In *SoutheastCon 2016* (pp. 1-6). IEEE.

Hopper, G. P. (1997). What Determines the Exchange Rate: Economic Factors or Market Sentiment. *Business Review*, 5, 17-29.

Huang, S. C., Chuang, P. J., Wu, C. F., & Lai, H. J. (2010). Chaos-based support vector regressions for exchange rate forecasting. *Expert Systems with Applications*, *37*(12), 8590-8598.

Ilyas, I. F., & Rekatsinas, T. (2022). Machine learning and data cleaning: Which serves the other? *ACM Journal of Data and Information Quality (JDIQ)*, *14*(3), 1-11.

Jamil, H. (2022). *Inflation forecasting using hybrid ARIMA-LSTM model* (Doctoral dissertation, Laurentian University of Sudbury).

Jovanovic, F., & Le Gall, P. (2001). Does God practice a random walk? The "financial physics" of a nineteenth-century forerunner, Jules Regnault. *European journal of the history of economic thought*, *8*(3), 332-362.

Jun Gu, W., hao Zhong, Y., zun Li, S., song Wei, C., ting Dong, L., yue Wang, Z., & Yan, C. (2024). Predicting stock prices with finbert-lstm: Integrating news sentiment analysis. In *Proceedings of the 2024 8th International Conference on Cloud and Big Data Computing* (pp. 67-72)

Kilian, L., & Taylor, M. P. (2003). Why is it so difficult to beat the random walk forecast of exchange rates? *Journal of International Economics*, *60*(1), 85-107.

Koru, G. K., & Uluyol, Ç. (2024). Detection of Turkish fake news from tweets with BERT models. *IEEE Access*, *12*, 14918-14931.

Lee, G. Y., Alzamil, L., Doskenov, B., & Termehchy, A. (2021). A survey on data cleaning methods for improved machine learning model performance. *arXiv preprint arXiv:2109.07127*.

Levy, D. (1994). Chaos theory and strategy: Theory, application, and managerial implications. *Strategic management journal*, *15*(S2), 167-178.

Lo, A. W. (1997). A Non-Random Walk Down Wall Street. In *Proceedings of symposia in pure mathematics* (Vol. 60, pp. 149-183).

Lo, A. W. (2004). The adaptive markets hypothesis: Market efficiency from an evolutionary perspective. *Journal of Portfolio Management, Forthcoming*.

Lo, A. W., & MacKinlay, A. C. (2011). *A non-random walk down Wall Street*. Princeton University Press.

Malkiel, B. G. (1999). *A random walk down Wall Street: including a life-cycle guide to personal investing*. WW Norton & Company.

Mareeswari, G., & Dinesh, E. V. (2023, March). Deep neural networks based

[DOI: 10.61882/jme.19.2.249]

detection and analysis of fake tweets. In *2023 4th International Conference on Signal Processing and Communication (ICSPC)* (pp. 56-61). IEEE.

Markova, M. (2019). Foreign exchange rate forecasting by artificial neural networks. In *AIP Conference Proceedings* (Vol. 2164, No. 1). AIP Publishing.

Meese, R. A., & Rogoff, K. (1983). Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of international economics*, *14*(1-2), 3-24.

Monica, C., & Nagarathna, N. J. S. C. S. (2020). Detection of fake tweets using sentiment analysis. *SN Computer Science*, *1*(2), 89.

Morina, F., Hysa, E., Ergün, U., Panait, M., & Voica, M. C. (2020). The Effect of Exchange Rate Volatility on Economic Growth: Case of the CEE Countries. *Journal of Risk and Financial Management, 13(8), 177.* doi:10.3390/jrfm13080177

Naeem, S., Mashwani, W. K., Ali, A., Uddin, M. I., Mahmoud, M., Jamal, F., & Chesneau, C. (2021). Machine learning-based USD/PKR exchange rate forecasting using sentiment analysis of Twitter data. *Computers, Materials & Continua*, *67*(3), 3451-3461.

Narayan, P. K. (2022). Understanding exchange rate shocks during COVID-19. *Finance Research Letters*, *45*, 102181.

Nazir, F., Ghazanfar, M. A., Maqsood, M., Aadil, F., Rho, S., & Mehmood, I. (2019). Social media signal detection using tweets volume, hashtag, and sentiment analysis. *Multimedia Tools and Applications*, *78*(3), 3553-3586.

Nikiforos, M. N., Vergis, S., Stylidou, A., Augoustis, N., Kermanidis, K. L., & Maragoudakis, M. (2020). Fake news detection regarding the Hong Kong events from tweets. In *IFIP International Conference on Artificial Intelligence Applications and Innovations* (pp. 177-186). Cham: Springer International Publishing.

Orellana, V., & Pino, G. (2021). Uncovered interest rate parity: A gravity-panel approach. *Heliyon*, 7(11).

Oussidi, A., & Elhassouny, A. (2018). Deep generative models: Survey. In *2018 International conference on intelligent systems and computer vision (ISCV)* (pp. 1-8). IEEE.

Ozcan, F. (2017). Exchange Rate Prediction from Twitter's Trending Topics. In Proceedings of the International Conference on Analytical and Computational Methods in Probability Theory, Moscow, Russia, (pp. 1–46).

Ozcelebi, O. (2018). Impacts of exchange rate volatility on macroeconomic and financial variables: Empirical evidence from PVAR modeling. In *Trade and Global Market*. IntechOpen.

Ozturk, S. S., & Ciftci, K. (2014). A sentiment analysis of twitter content as a predictor of exchange rate movements. *Review of Economic Analysis*, *6*(2), 132-140.

Pandey, T. N., Jagadev, A. K., Dehuri, S., & Cho, S. B. (2018). A review and empirical analysis of neural networks-based exchange rate prediction. *Intelligent Decision Technologies*, 12(4), 423-439.

[DOI: 10.61882/jme.19.2.249]

Pandey, T. N., Jagadev, A. K., Dehuri, S., & Cho, S. B. (2020). A novel committee machine and reviews of neural network and statistical models for currency exchange rate prediction: An experimental analysis. *Journal of King Saud University-Computer and Information Sciences*, *32*(9), 987-999.

Patton, D. U., Frey, W. R., McGregor, K. A., Lee, F. T., McKeown, K., & Moss, E. (2020). Contextual analysis of social media: The promise and challenge of eliciting context in social media posts with natural language processing. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (pp. 337-342).

Qi, B., Costin, A., & Jia, M. (2020). A framework with efficient extraction and analysis of Twitter data for evaluating public opinions on transportation services. *Travel behaviour and society*, *21*, 10-23.

Rossi, B. (2013). Exchange rate predictability. *Journal of economic literature*, *51*(4), 1063-1119.

Schröer, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. *Procedia Computer Science*, *181*, 526-534.

Serrano-Guerrero, J., Olivas, J. A., Romero, F. P., & Herrera-Viedma, E. (2015). Sentiment analysis: A review and comparative analysis of web services. *Information Sciences*, *311*, 18-38.

Shah, P., Desai, K., Hada, M., Parikh, P., Champaneria, M., Panchal, D., ... & Shah, M. (2024). A comprehensive review on sentiment analysis of social/web media big data for stock market prediction. *International Journal of System Assurance Engineering and Management*, *15*(6), 2011-2018

Sharma, S., Saraswat, M., & Dubey, A. K. (2022). Fake news detection on Twitter. *International Journal of Web Information Systems*, *18*(5/6), 388-412.

Shelke, N. M., Deshpande, S., & Thakre, V. (2012). Survey of techniques for opinion mining. *International Journal of Computer Applications*, 57(13).

Shiller, R. J. (2017). *Narrative Economics. American Economic Review, 107(4), 967–1004.* doi:10.1257/aer.107.4.967

Shiller, R. J. (2019). Narrative economics: How stories go viral and drive major economic events. *The Quarterly Journal of Austrian Economics*, *22*(4), 620-627.

Shui-Ling, Y. U., & Li, Z. (2017). Stock price prediction based on ARIMA-RNN combined model. In *4th International Conference on Social Science (ICSS 2017)* (pp. 1-6).

Sidehabi, S. W., & Tandungan, S. (2016). Statistical and machine learning approach in forex prediction based on empirical data. In *2016 International Conference on Computational Intelligence and Cybernetics* (pp. 63-68). IEEE.

Sirignano, J., & Cont, R. (2019). Universal features of price formation in financial markets: perspectives from deep learning. *Quantitative Finance*, *19*(9), 1449-1459.

Sonia, J. J., Goenka, A., & Jain, A. (2024). Stock price analysis using sentiment analysis of twitter data. In *2024 International Conference on Intelligent Systems for Cybersecurity (ISCS)* (pp. 01-07). IEEE.

[DOI: 10.61882/jme.19.2.249]

Susan, S., & Kumar, A. (2021). The balancing trick: Optimized sampling of imbalanced datasets - A brief survey of the recent State of the Art. *Engineering Reports*, *3*(4), e12298.

Taylor, M. P., & Allen, H. (1992). The use of technical analysis in the foreign exchange market. *Journal of international Money and Finance*, *11*(3), 304-314.

Temür, A. S., Akgün, M., & Temür, G. (2019). Predicting housing sales in Turkey using ARIMA, LSTM and hybrid models. *Journal of business economics and management*, *20*(5), 920-938.

Vishwakarma, A., Singh, A., Mahadik, A., & Pradhan, R. (2020). Stock price prediction using Sarima and Prophet machine learning model. *Journal of Advanced Research in Science, Communication and Technology (IJARSCT)*, *9*(1).

Wang, R., Morley, B., & Stamatogiannis, M. P. (2019). Forecasting the exchange rate using nonlinear Taylor rule based models. *International Journal of Forecasting*, *35*(2), 429-442.

White, M. D., & Marsh, E. E. (2006). Content analysis: A flexible Methodology. *LibraryTrends, 55 (1), 22-45*.

Wirth, R., & Hipp, J. (2000, April). CRISP-DM: Towards a standard process model for data mining. In *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining* (Vol. 1, pp. 29-39).

Xu, Z. (2003). Purchasing power parity, price indices, and exchange rate forcasts. *Journal of International Money and Finance*, *22*(1), 105-130.

Yasir, M., Durrani, M. Y., Afzal, S., Maqsood, M., Aadil, F., Mehmood, I., & Rho, S. (2019). An intelligent event-sentiment-based daily foreign exchange rate forecasting system. *Applied Sciences*, *9*(15), 2980.

Appendix

Appendix 1: "Daily Dollar Sentiment" feature construction Method

Performing the sentiment analysis model inference step leads to determining the probability of dependency for each tweet (a real number within the range [0,1]) to each of the three classes: "Positive", "Negative", and "Neutral" at the softmax layer of the ParsBERT model (Farahani et al., 2021). The algorithm used to determine the sentiment class for each tweet selects the class with the highest probability among the others. Assuming that the numerical label for the three classes is, respectively, +1, -1 and 0, then one of the common methods for calculating the quantitative value of sentiment for a day is obtained using the following formula:

[DOI: 10.61882/jme.19.2.249]

$$Daily\ Numerical\ Sentiment\ Value = \frac{(\#\ of\ POS\ tweets\ per\ day \times 1) + (\#\ of\ NEG\ tweets\ per\ day \times -1) + (\#\ of\ NEU\ tweets\ per\ day \times 0)}{Total\ number\ of\ tweets\ per\ day}$$

The values obtained from the above formula for each day are real numbers in the range [-1,1]. A drawback of this computational method is the elimination of the weight of neutral tweets (by applying a zero coefficient), while their weight is accounted for differently in the denominator. Therefore, to map the obtained value to one of the three sentiment classes for each day, the following rules were employed, using coefficients α and β (Table 19).

Table 19
Rules of membership in daily sentiment classes

Class Name	Class Numerical label	Class Membership Rule
Positive sentiment	+1	The Daily Numerical Sentiment Value is greater than the threshold α+β.
Negative sentiment	-1	The Daily Numerical Sentiment Value is less than the threshold α-β.
Neutral sentiment	0	The Daily Numerical Sentiment Value is between or equal to either of the thresholds α+β and α-β.

Source: Research findings

In Figure 9, the position and logical relationship of parameters α and β with the permissible values of the "Daily Dollar Sentiment" feature are illustrated.

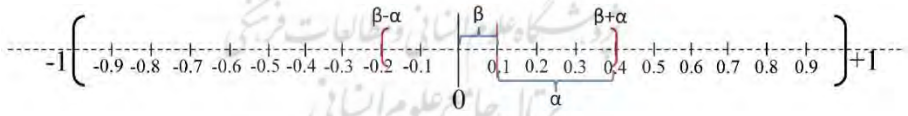

Figure 9. A schema of the range of "numerical value of daily dollar sentiment" and the position of values for α and β parameters in it
Source: Research findings

Observations indicated that Iranian users' initial reaction to a shift in market rate direction (which generally trends upward) from "increase" to "decrease" or "neutral" is typically cautious and inclined to believe in the continuation of the stabilized previous trend. From a broader perspective, in a market where the rate direction remains stable for most periods, users are expected to respond to a directional change from a stabilized state with caution and a tendency to believe in the persistence of the stabilized trend. Therefore,

[DOI: 10.61882/jme.19.2.249]

since the exchange rate in the Iranian currency market has generally exhibited an increasing trend over a relatively long period (several years), this pattern has remained in the memory of the people and market players. Consequently, the neutral point does not start from zero but instead begins at another initial positive value. The parameter β determines the distance of the neutral baseline from zero in the studied market, while the parameter α determines the range of neutral class values and its boundary with the other two classes. To select the most appropriate values for these two parameters, their optimal values were determined at the point of maximum correlation between the numerical value of the "Daily Dollar Sentiment" feature and the actual daily exchange rate fluctuation. The values obtained for the α and β parameters were 0.3 and 0.1, respectively.

Appendix 2: Description of data interpolation and extrapolation process

Observations indicated that some time series contained missing data, often occurring frequently but primarily over short time intervals. These gaps were generally due to local market closures on weekends, the higher-frequency publication schedule than daily, or, in the worst cases, the data source not publishing any data. Additionally, for some variables, a similar phenomenon was observed not within the time series but at its endpoints. To address and reconstruct the missing data, interpolation and extrapolation methods were applied sequentially. Various techniques exist for these tasks, including:

1) Rule-based methods, which may have significant deviations when the time series gaps are large (e.g., interpolating missing values in datasets with seasonal publication frequencies). These methods were not used in this research.
2) Estimation-based methods, which leverage predictive approaches as a key advantage. These methods typically require robust assumptions that are not readily available in real-world, and were therefore not utilized in this study.
3) Mathematical modeling methods, both linear and nonlinear. While nonlinear models offer greater flexibility, they often require substantial processing power. Additionally, some of nonlinear models often need access to four points close to the gap in the time series to perform data recovery, which is often impractical due to data unavailability.

Given that most macroeconomic variables exhibit a relatively stable nature, with general trends following a linear pattern and fluctuations becoming apparent over the long term, reconstructing missing in-sample data

[DOI: 10.61882/jme.19.2.249]

using simple linear algorithms will not distort data patterns over short intervals. Therefore, in this study, linear interpolation was employed for this purpose. Figure 10 illustrates the interpolation process for the variable "US Treasury (5Y)", where the reconstructed data points are highlighted in red.

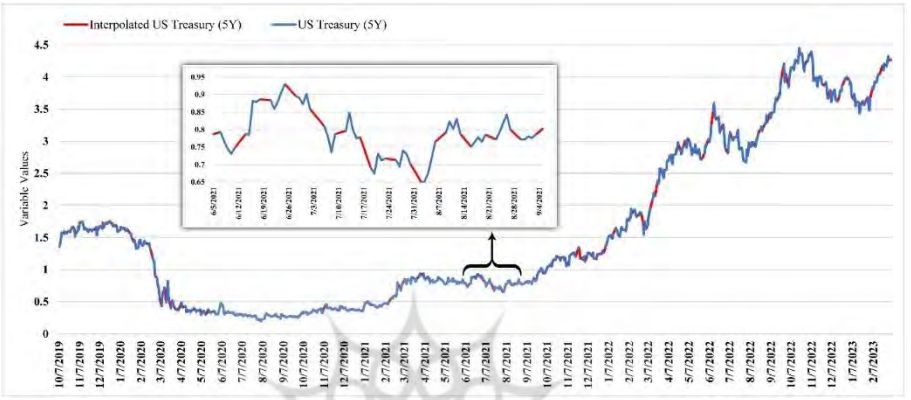

Figure 10. A view of the interpolated time series data for the variable "US Treasury (5Y)"
Source: Research findings

In the process of data extrapolation, unlike interpolation (where data on both sides of the missing gap is available), only the lower boundary value is accessible. Therefore, to accurately reconstruct the missing data from the lower boundary to the end of the series, it is necessary to use time series forecasting approaches. Research has shown that using linear econometric forecasting methods for time series can provide reasonable estimates of future values. Among linear forecasting methods for time series, the ARIMA (Autoregressive Integrated Moving Average) models, are widely used and have demonstrated excellent performance. A variant of this method, called SARIMA, extends ARIMA by incorporating seasonal component into the model. SARIMA has shown superior performance when dealing with univariate time series data that exhibits seasonal shocks (Vishwakarma et al., 2020). Thus, the present study adopted this model to extrapolate some of the missing data. To determine the optimal SARIMA model for each variable, one approach is to evaluate the model's performance across all possible parameter values. However, since this process requires substantial processing resources and is time-consuming, a constrained search space was adopted in this study. For all AR, I, and MA components in the model structure $(p, d, q)(P, D, Q)_s$,

[DOI: 10.61882/jme.19.2.249]

integer values from 0 to 2 were considered for each of the six main parameters, resulting in 729 possible combinations. Additionally, seasonal frequencies of 7, 14, 21, and 28 were considered for the seasonal parameter, leading to a total of 2,916 different SARIMA model to be tested for each variable. Once each of the 2916 models was constructed for the 11 variables requiring extrapolation, model performance was evaluated using the Akaike Criteria (AIC). The model with the lowest AIC value was selected as the best fit for predicting each variable, and ultimately, the missing data were extrapolated using the chosen model. In Figure 11, the red section represents the extrapolated data for the PMI variable.

Figure 11. A view of the extrapolated time series data for the variable "Purchasing Managers Index (PMI)"
Source: Research findings

Appendix 3: Sources of economic data

The data for economic variables were gathered from various financial data sources in Iran and the USA, as well as world's financial data providers. Table 20 presents the details of these sources, along with the number of variables provided by each source.

[DOI: 10.61882/jme.19.2.249]

Table 20
Data sources used for collecting economic variables

No.	Title	Short description	Source Location	No. of variables
1	Federal Reserve Economic Data (FRED)	A US organization responsible for leading financial and commodity exchange markets.	USA	4
2	Intercontinental Exchange (ICE)	An American company responsible for managing financial and commodity exchange marketplaces.	USA	1
3	Yahoo Finance	The economic section of Yahoo's website, which publishes economic data.	USA	6
4	Central Bank of the Islamic Republic of Iran (CBI)	The statistics and data section on the website of the Central Bank of Iran.	Iran	2
5	Statistical Centre of Iran (SCI)	The National Statistics Portal of the Statistical Center of Iran.	Iran	4
6	Iran Chamber Research Center	The Research Center of the Chamber of Commerce, Industry, Mines, and Agriculture.	Iran	1
7	Tehran Securities Exchange Technology Management Co.	The technical infrastructure management company of the Tehran Stock Exchange.	Iran	2
8	BourseView	A reputable source for regularly releasing economic variables such as the US dollar exchange rate and gold coin prices.	Iran	2
9	Gold and Currency Information Network	The website of the Gold and Jewelry Union, which publishes rates.	Iran	2

Source: Research findings

Appendix 4: Fake tweet reviewing, identification, and removal process

In the first step, during the review of related literature, several general heuristic and rules of thumb describing qualitative characteristics of fake tweets were identified.

In the second step, to identify suspiciously fake tweets, three quantitative rules were formulated based on the qualitative rules obtained in the previous step and the available data for each tweet. These rules were implemented in a computer program, which after execution, ultimately identified and extracted a total of 442,370 suspicious tweets.

[DOI: 10.61882/jme.19.2.249]

In the third step, various samples, patterns, and categories of fake tweets were identified, analyzed, and classified through human and statistical review and analysis of the extracted tweets. For instance, tweets with similar content were categorized into three main groups, with some of these groups further divided into subcategories.

In the fourth step, appropriate actions were determined for each identified category and group based on the conducted analyses. Some actions were implemented using rules of thumb (formulated in this step) and applied to suspicious tweets through a computer program. Other actions were carried out using a combination of human intervention (where certain tweets were manually labeled as exceptions to be excluded from the automated rules) and machine-based processing (Figure 12).

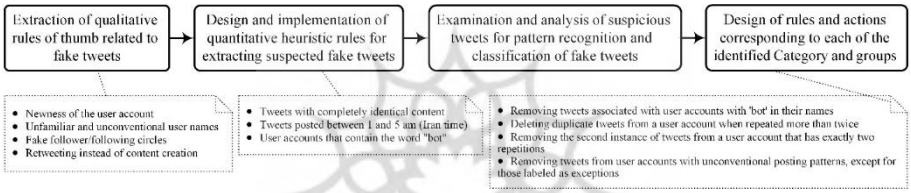

Figure 12. Main steps in the fake tweet identification and removal process
Source: Research findings

Table 21 presents the heuristic rules for removing fake tweets and their underlying reasons for each. Some of the rules overlap in scope; however, during the implementation phase, any overlapping cases were resolved according to the order specified in the table, ensuring that they were accounted for in the first applicable rule.

[DOI: 10.61882/jme.19.2.249]

Table 21
Rules of thumb designed for removing fake tweets and their rationale

No.	Rule	Reason for action
1	Removing tweets associated with user accounts with 'bot' in their names	Such accounts are likely non-human agents, which raises suspicion about the true intent behind the expressed sentiment in their tweets. In other words, the sentiment expressed in their tweets may serve purposes other than personal opinion.
2	Deleting duplicate tweets from a user account when repeated more than twice	The excessive and unusual repetition of a tweet by a single account may indicate an attempt to manipulate market sentiment or public opinion, so the sentiment in such tweets can be misleading or unreliable.
3	Removing the second instance of tweets from a user account that has exactly two repetitions	In many cases, the second duplicate tweet results from user error, particularly among Iranian users who rely on auxiliary tools such as VPNs to access X, which can reduce the network's efficiency.
4	Removing tweets from user accounts with unconventional posting patterns, except for those labeled as exceptions	These accounts typically belong to news agencies, stock traders, currency market businesses, and automated agents focusing on disseminating information rather than expressing personal sentiments, making their tweets unreliable for sentiment analysis.

Source: Research findings

[DOI: 10.61882/jme.19.2.249]