Sentimental Analysis to Predict Stock Market Using in Neutrosophic Time Series

Saravanaraj .S .S[1], Vediyappan Govindan[1,*], Said Broumi[2], Haewon Byeon[3,*]

[1]Department of Mathematics, Hindustan Institute of Technology and Science, Chennai, Tamil Nadu 603103, India
[2]Laboratory of Information Processing, Faculty of Science Ben M'Sik, University, Hassan II, B.P 7955, Morocco
[3]Department of AI Big data, Inje University, Gimhae, 50834, Republic of Korea

Emails: saravanarajcibi2000@gmail.com; govindoviya@gmail.com; broumisaid78@gmail.com; byeon@inje.ac.kr

Abstract

This study delves into the innovative use of sentiment analysis in conjunction with neutrosophic time series to forecast stock market trends in various contexts. By meticulously analyzing financial news and social media data, sentiment scores are derived and subsequently integrated into a neutrosophic time series model. This model is uniquely adept at managing uncertainty and indeterminacy, providing a robust framework for prediction. The findings indicate that this integrated approach significantly enhances predictive accuracy and reliability over traditional time series models. This research presents a novel methodology for tackling the intrinsic unpredictability of stock markets, offering a more reliable tool for investors and analysts across diverse financial environments. Additionally, by incorporating sentiment scores from a wide range of sources, the model captures a comprehensive view of market sentiment, reflecting the collective mood and opinions of investors. This comprehensive approach ensures that the predictions are not only accurate but also reflective of real-time market dynamics. Finally, this work highlights the possibility of merging sentiment analysis with sophisticated modeling approaches to change stock market prediction, as well as providing a promising avenue for future financial forecasting research.

Keywords: Stock Market Prediction; Neutrosophic Time Series; Sentiment Analysis

1. Introduction

Forecasting patterns in the stock market is critical, but difficult because of the inherent volatility and psychological forces that drive market behavior. Traditional methods of stock market prediction primarily focus on quantitative data, such as historical prices, trading volumes, and economic indicators. While these methods have their merits, they often overlook the emotional and psychological factors that can significantly affect market dynamics. Sentiment analysis, which extracts emotions and opinions from textual data sources like financial news and social media, offers a valuable new dimension to market analysis. By capturing public sentiment, these methods provide insights into the behavioural aspects of market participants, which are not reflected in purely quantitative data.

Mathematician and philosopher Florentin Smarandache (1) created neoclassical logic in the late 20th century, which is a substantial development of conventional logic. Unlike classical logic, which operates within binary frameworks of true and false, neutrosophic logic introduces the concepts of truth, falsehood, and indeterminacy to handle uncertainty and contradictions. This logic system is very useful when dealing with ambiguous, imprecise, or contradictory information, so it is suitable for a variety of situations in life where traditional reasoning falls short.

DOI: https://doi.org/10.54216/IJNS.250215
Received: February 14, 2024 Revised: May 02, 2024 Accepted: August 08, 2024

By embracing the degrees of reality, falsehood, and uncertainty, neutrosophic logic gives a more nuanced and flexible approach to reasoning, making it more effective in disciplines that involve artificial intelligence, decision-making processes, and data analytics.

However, sentiment analysis introduces its own set of challenges, particularly related to the uncertainty and indeterminacy of textual information. To effectively address these challenges, Kandasamy.I (2) introduced sentiment analysis within the framework of a collection of neutrosophic elements. We are using to combine the sentiment analysis and neutrosophic time series to predict the stock market, it is a mathematical framework designed to handle incomplete, indeterminate, and uncertain information using neutrosophic logic. Neutrosophic logic extends traditional fuzzy logic by allowing the representation of truth, falsehood, and indeterminacy simultaneously, making it well suited for the complexities of sentiment-laden data.

This study seeks to improve the correctness and dependability of forecasts of the stock market by incorporating sentiment ratings from financial news and social media into a neutrosophic time series model. We hypothesize that this integrated approach will outperform traditional time series models, providing a more robust tool for stock market analysis. The subsequent sections of this paper review relevant literature, detail our methodology, present and discuss the experimental results, and conclude with implications for future research.

2. Related Works

In 2019, Abdel-Basset et al., (3) made significant strides in improving forecasting accuracy rates compared to existing methodologies such as fuzzy and intuitionistic time series. They introduced initial and advanced neutrosophic time series (NTS) data, which enhanced the precision of forecasting techniques. A key step in the neutrosophication process involves obtaining triangular neutrosophic numbers, essential for representing historical time series data within a neutrosophic framework, capturing uncertainty, indeterminacy, and inconsistency. This allows for a richer capture of the complex characteristics of time series data. To convert neutrosophic data back into a conventional form, the deneutrosophication procedure is employed, involving straightforward arithmetic calculations for both initial and advanced NTS. This process translates enriched neutrosophic data into a format suitable for traditional analysis, ensuring the benefits of the neutrosophic framework are retained while making the data accessible for further analysis. By leveraging these innovative approaches, Abdel-Basset et al. provided a robust framework for more accurate and reliable time series forecasting, paving the way for advancements in predictive analytics and offering significant promise for applications such as financial forecasting and weather prediction.

Hongjun Guan (4) proposed a unique forecasting framework in 2019 that is based on Neutrosophic Set Theory and information entropy, with a focus on high-order fuzzy fluctuation data. This approach, known as Neutrosophic Fuzzy Time Series (NFTS), uses neutrosophic set theory to improve time series data analysis by measuring the complexity of historical fluctuations. Guan's approach involves identifying Neutrosophic Logical Relationships (NLRs) using information entropy and NFTS, leading to the development of a comprehensive structure for evaluating and simulating the uncertainty in historical trends. The model's efficacy was demonstrated through its application to three common test datasets, highlighting its stability and universality across different scenarios. This innovative method improves time series forecasting by effectively capturing and analyzing the complex and uncertain nature of historical data, providing a robust tool for more accurate and reliable predictions in various applications.

In 2019, Pritpal Singh (5) introduced an innovative hybrid framework for forecasting time series that combines the gradient descent approach, neural network prediction (ANN), and neutrosophic set (NS) theory. This innovative approach begins with creating an ANN architecture to perform the deneutrosophication process from Neutrosophic Entropy Difference Relationships (NEDRs). The architecture is trained using the Backpropagation Neural Network (BPNN) method, which employs the gradient descent algorithm to optimize the Average Forecasting Error Rate (AFER). By iteratively adjusting the weights within the neural network, the BPNN method aims to minimize discrepancies between the target and computed output entropy levels, enhancing prediction accuracy. Integrating neutrosophic set theory provides a more thorough picture of uncertainty, indeterminacy, and inconsistency within the data, addressing ambiguities that traditional methods may overlook. This hybrid model's ability to capture and analyze complex patterns makes it valuable for applications such as financial forecasting and weather prediction. Singh's work demonstrates the model's robustness and versatility, significantly advancing time series analysis and offering a sophisticated method for forecasting in uncertain and dynamic environments

Aiwu Zhao (6) created a multi-attribute fuzzy oscillation time series-forecasting model in 2019 by fusing neutrosophic soft sets (NSSs) with information entropy. The model takes use of NSSs' capacity to describe several dimensions and features, accurately portraying the stock market's complicated rules and volatility. The

DOI: https://doi.org/10.54216/IJNS.250215
Received: February 14, 2024 Revised: May 02, 2024 Accepted: August 08, 2024

inconsistency and unpredictability of market movements are measured using information entropy, which provides a quantitative measurement of their volatility at any given time. This integrated approach offers a comprehensive analysis by considering various aspects of market behavior, making it more nuanced than traditional forecasting methods. The empirical findings demonstrate the model's accuracy in predicting stock market trends, highlighting its potential as a valuable tool for financial analysis and decision-making. By explaining difficult modifications throughout the logic pattern-training step, the model demonstrates a strong ability to adapt to changing market conditions, providing a comprehensive framework for understanding and forecasting market dynamics.

In 2020, P. Singh (7) created an enhanced time- series forecasting technique that integrates the particle swarm optimization (PSO) algorithm with neutrosophic set (NS) theory, yielding a revolutionary method that tackles the inherent uncertainty and imprecision in time series data. The model utilizes the Neutrosophic Time Series (NTS) technique to effectively represent and analyze the data, leveraging NS theory's ability to manage incomplete and indeterminate information. The PSO algorithm enhances the model's predictive accuracy by optimizing its parameters, enabling the model to better capture underlying patterns and trends. Comparative analysis with other forecasting models demonstrates the superior performance of the NTS-PSO model, highlighting its potential for real-world applications in fields like finance, economics, and engineering. Despite its strengths, the study identifies a limitation with univariate datasets, suggesting that future research could explore the model's effectiveness with multivariate data to further improve its robustness and applicability. This study marks a substantial development in time series forecasting by providing a sophisticated tool for making reliable forecasts in complicated data contexts.

In 2024, Seyyed Ahmad Edalatpanah et al. (8) improve the performance of the NTS modeling approach by employing the quantum optimization algorithm (QOA), the genetic algorithm (GA), and particle swarm optimization. These optimization techniques use a set of locally optimum solutions to determine the speech's global context and related intervals. The suggested hybrid model, the NTS-QOA model, is verified using datasets from Alabama university enrollment, the Taiwan futures exchange (TAIFEX) index, and the Taiwan Stock Exchange Corporation (TSEC) weighted index. The average forecasting error rate (AFER) shows that the NTS-QOA model outperforms existing benchmark models in terms of efficiency. The NTS-QOA, NTS-GA, and NTS-PSO algorithms' respective AFER values for the university dataset are 0.166, 0.167, and 0.164. The AFER for the TAIFEX dataset is 0.081, while for the TSEC dataset, it is 0.09. The prediction results show that the suggested models outperform existing ones, with the NTS-PSO model outperforming the NTS-QOA and NTS-GA models in terms of convergence rate and reduced AFER. This strategy, which incorporates neutrosophic logic and optimization approaches, improves predicting accuracy greatly.

3. Methodology

The methodology for this study involves several key steps, including data collection, sentiment analysis, the formulation of a neutrosophic time series model, and the integration of sentiment scores into this model. Here is a detailed explanation.

3.1 Data Collection

We are utilizing the StockNet benchmark dataset (9). This dataset includes tweets as well as historical stock market data for many stocks. The period from December 31, 2014 to December 31, 2015. This dataset is useful for applications like using sentiment analysis of financial news and news sentiment to anticipate stock prices, event studies on the influence of news events, and temporal research of the effects of news on stock prices. The tweets were gathered from Twitter using NASDAQ symbols such as AAPL (Apple). When historical data for the same time was gathered from Yahoo Finance. Preprocessing these datasets includes normalizing stock prices, resampling to a consistent frequency, managing missing values and transforming text data into sentiment scores using techniques like tokenization and vectorization. Economic indicators are normalized and missing values imputed. All data is temporally aligned and augmented with features like lagged values and moving averages. Finally, the data is translated into a neutrosophic format, in which each point is represented by degrees of reality, uncertainty, and falsehood, in order to accurately simulate the financial market's unpredictability and complexity.

3.2 Sentiment Analysis

The pre-processed text is then quantitatively expressed utilizing methods such as TF-IDF, Word2Vec, or BERT embedding. The text is assigned sentiment values ranging from binary (positive/negative) to continuous scales (for example, -1 to 1). These sentiment scores are aligned with historical stock data and economic indicators to form a comprehensive dataset that reflects market sentiment, enabling more informed predictions of stock market trends.

DOI: https://doi.org/10.54216/IJNS.250215
Received: February 14, 2024 Revised: May 02, 2024 Accepted: August 08, 2024

3.3 Neutrosophic Time Series Model

In sentimental analysis, we are forecasting the stock market using the neutrosophic time series model. In the explanation of the method in Figure 1.

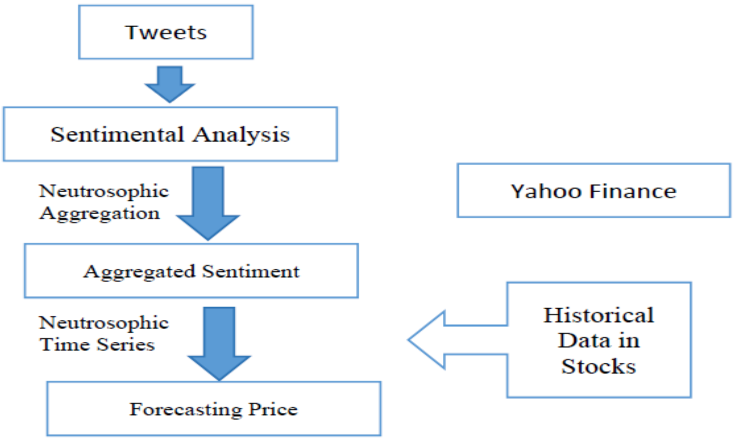

Figure1. Sentimental analysis in Neutrosophic time series

In the algorithm of Neutrosophic, they give time series

Step1: Using the data at hand, define the universe of discourse (U). Choose the greatest d_l and the d_s lowest from the provided data, d_v then

U $= [d_s - d_1, d_l + d_2]$

In d_1 & $d_2$do issue domain experts provide two relevant positive numbers.

Step2: Divide the world of discussion into m triangular neutrosophic numbers

$$m = \frac{d_l + d_2 - d_s - d_1}{le}$$

Step3: Build the triangular neutrosophic numbers based about the number of triangular neutrosophic numbers on the discourse universe and specified length (*le*). The triangular neutrosophic numbers is $\widetilde{N_1}, \widetilde{N_2}, \cdots \widetilde{N_m}$ .In the triangular neutrosophic numbers in $\widetilde{N} = \langle(\eta_1, \eta_2, \eta_3); T_{\widetilde{N}}, I_{\widetilde{N}}, F_{\widetilde{N}}\rangle$ is a special neutrosophic set on Real number R. In the truth membership, indeterminate membership, and non-membership functions. In the degree of memberships as in (10)

$$T_{\widetilde{N}}(x) = \begin{cases} T_{\widetilde{N}}\left(\dfrac{x-\eta_1}{\eta_2-\eta_1}\right) & (\eta_1 \le x \le \eta_2) \\ T_{\widetilde{N}} & (x = n_2) \\ T_{\widetilde{N}}\left(\dfrac{\eta_1 - x}{\eta_3-\eta_2}\right) & (\eta_2 < x \le \eta_3) \\ 0 & otherwise, \end{cases}$$

$$I_{\widetilde{N}}(x) = \begin{cases} \left(\dfrac{\eta_2 - x + I_{\widetilde{N}}(x-\eta_1)}{\eta_2-\eta_1}\right) & (\eta_1 \le x \le \eta_2) \\ I_{\widetilde{N}} & (x = \eta_2) \\ \left(\dfrac{x - \eta_2 + I_{\widetilde{N}}(\eta_1 - x)}{\eta_3-\eta_2}\right) & (\eta_2 < x \le \eta_3) \\ 1 & otherwise, \end{cases}$$

$$F_{\widetilde{N}}(x) = \begin{cases} \left(\dfrac{\eta_2 - x + F_{\widetilde{N}}(x-\eta_1)}{\eta_2-\eta_1}\right) & (\eta_1 \le x \le \eta_2) \\ F_{\widetilde{N}} & (x = \eta_2) \\ \left(\dfrac{x - \eta_2 + I_{\widetilde{N}}(\eta_1 - x)}{\eta_3-\eta_2}\right) & (\eta_2 < x \le \eta_3) \\ 1 & otherwise, \end{cases}$$

DOI: https://doi.org/10.54216/IJNS.250215
Received: February 14, 2024 Revised: May 02, 2024 Accepted: August 08, 2024

Step 4: Construct a neutrosophication process from the existing data. If $i, j = 1,2,3 \cdots v$ (the end of data):

Rule 1: To compute the score degree, use this equation:

$$SC_{\widetilde{N_j}}(x_i) = 2 + T_{\widetilde{N}_j}x_i - I_{\widetilde{N}_j}x_i - F(x_i)$$

Rule 2: Utilize the equation that follows to determine the board level and choose the most inaccurate degree if two neutrosophic numbers have a comparable board level:

$$AC_{\widetilde{N_j}}(x_i) = 2 + T_{\widetilde{N}_j}x_i - I_{\widetilde{N}_j}x_i + F(x_i)$$

Step 5: Create the following neutrosophic logical relationships (NLRs):

In the situation that $\widetilde{N}_j, \widetilde{N_k}$ represent the neutrosophication quantities for years n and n + 1, correspondingly, then $\widetilde{N}_j \rightarrow \widetilde{N_k}$ represents the NLR.

Step 6: In order to assess the neutrosophic logical relationship groups (NLRGs), identify variables and their neutrosophic sets, construct and group neutrosophic relationships, formulate NLRGs, and validate them using a subset of the data.

Step 7: The projected value in this scenario will be (i.e., left empty) if the converted the crisp value to neutrosophic value of data $_i$ is $\widetilde{N_k}$ it is unaffected by any other converted the crisp value to neutrosophic values, and analysing the NLRG for this value finds no dependencies (i.e., $\neq \rightarrow \widetilde{N_k}$), The $\neq$, sign indicates no value.

Step 8: To compute the forecasting error, use the formulae below:

Root Mean Square Error (RMSE) = $\sqrt{\frac{\sum_{i=1}^{n}(Forecast_i - Actual_i)^2}{n}}$

Forecasting Error = $\frac{|Forecast - Actual|}{Actual} \times 100$

4. Results and discussions

Sentiment scores were integrated with stock price time series data to create the Neutrosophic time series model. The model parameters were optimized to account for the influence of each sentiment component on the stock prices. The neutrosophic model was then used to forecast future stock prices, and its performance was evaluated against actual stock market data. In the neutrosophic time series sentimental analysis in figure 3.

4.1 Comparative Analysis with Fuzzy Time Series

We also used the same emotion data to create a fuzzy time series model for comparison. The fuzzy time series model used membership functions to manage sentiment score ambiguity, but it did not explicitly account for the indeterminacy component. The same stock market data was used to test the fuzzy time series model's performance. Figure 2 depicts the fuzzy time series sentimental analysis in (11).

1. Prediction Accuracy: The neutrosophic time series model demonstrated higher prediction accuracy compared to the fuzzy time series model. By incorporating the indeterminacy component, the neutrosophic model was able to capture the inherent uncertainty in the stock market data more effectively.
2. Handling of Indeterminacy: The neutrosophic model's ability to explicitly handle indeterminacy provided a significant advantage in scenarios where market sentiments were ambiguous or contradictory. This led to more robust predictions, especially during periods of market volatility.
3. Computational Complexity: The neutrosophic time series model was computationally more demanding than the fuzzy time series model. The need to process and integrate the three components (truth, falsity, and indeterminacy) resulted in increased processing time and resource requirements.
4. Interpretability: While the neutrosophic model offered a nuanced approach to handling uncertainty, it was also more complex to interpret. The fuzzy time series model, with its simplified structure, was easier to grasp and apply, providing more feasible for some applications.

The findings of this study emphasize the potential of neutrosophic time series in improving stock market forecast using sentiment analysis. By explicitly incorporating the indeterminacy component, the neutrosophic model gives an extra thorough technique for uncertainties handling. This is particularly useful in financial markets, where investor sentiment can be highly unpredictable and subject to rapid changes.

DOI: https://doi.org/10.54216/IJNS.250215
Received: February 14, 2024 Revised: May 02, 2024 Accepted: August 08, 2024

However, the increased computational complexity and difficulty in interpretation present challenges that need to be addressed. We will compare Neutrosophic and fuzzy time series in table1. Neutrosophic time series offer the strength of handling indeterminacy by considering truth, indeterminacy, and falsity, providing a advanced approach to uncertainty. However, they can be complex and computationally demanding. On the other hand, fuzzy time series are simpler and effective in managing vagueness in data, making them easier to implement and interpret. However, they lack the explicit handling of indeterminacy, which can sometimes limit their precision in capturing complex uncertainties.

Table1: Compare the three time series models.

Models	RMSE	Average Forecasting Error Percentage (%)
Time series model	0.1567	54.892%
Fuzzy time series	0.1140	63.818%
Neutrosophic time series	0.1036	72.818%

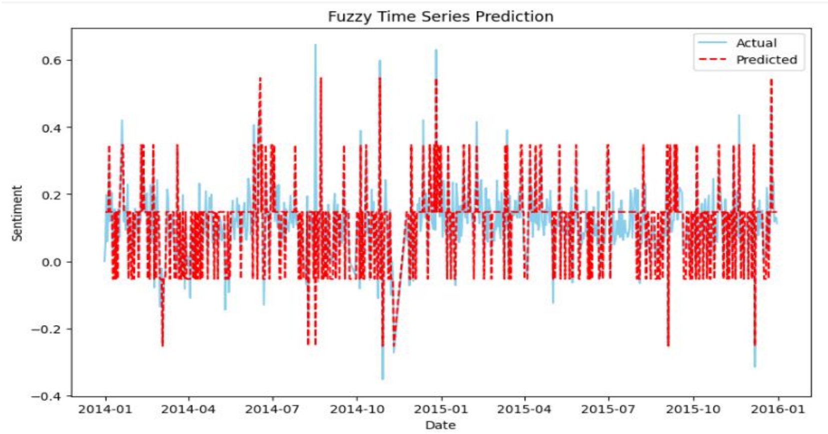

Figure 2. Fuzzy time series

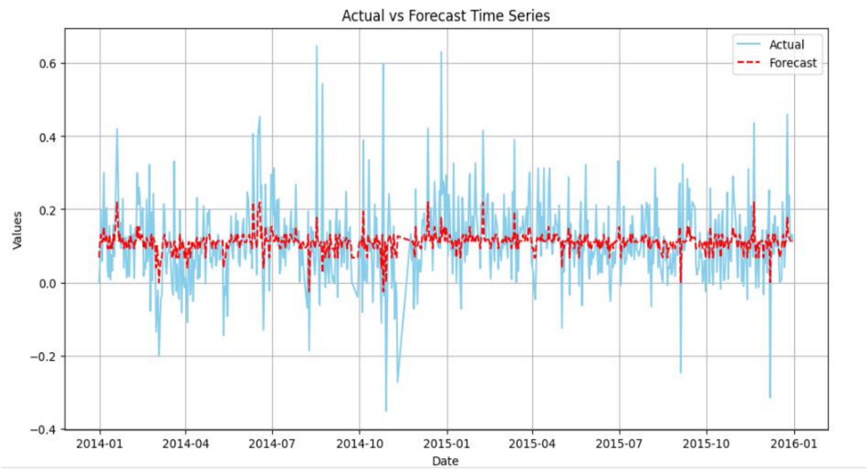

Figure 3. Neutrosophic time series

5. Conclusion and future works

Furthermore, the neutrosophic time series model provided a sophisticated framework for managing the complexities and uncertainties inherent in the financial markets. By effectively differentiating between certain (truth), uncertain (indeterminacy), and false (falsity) information, the model allowed for a more refined analysis of market data. This capability enabled more robust and reliable predictions, even under volatile market conditions, where traditional models often struggle to maintain accuracy. The abilities of the model to integrate and handle a wide range of data—from quantitative financial measurements to qualitative sentiment data—distinguishes it from

DOI: https://doi.org/10.54216/IJNS.250215
Received: February 14, 2024 Revised: May 02, 2024 Accepted: August 08, 2024

traditional techniques, which may neglect or insufficiently address qualitative components of market sentiment.In particular, the neutrosophic time series approach excels in scenarios where market sentiment is ambiguous or contradictory. Traditional models may fail to capture these nuances, leading to oversimplified or inaccurate predictions. In contrast, our approach leverages the multi-dimensional nature of neutrosophic logic to provide a more comprehensive view of market dynamics. This is especially valuable in today's fast-paced and information-rich financial landscape, where investor sentiment can change rapidly and unpredictably.

Overall, the study underscores the significant potential of combining sentiment analysis with advanced mathematical frameworks like neutrosophic logic to enhance stock market prediction. This integrated approach not only captures the strengths of sentiment analysis in gauging public opinion but also leverages the sophisticated uncertainty management capabilities of neutrosophic time series. By taking into account the complete range of reality, uncertainty, and falsehood, the model provides a more comprehensive and accurate picture of market circumstances.

Funding: This research Supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF- RS-2023-00237287, NRF-2021S1A5A8062526) and local government-university cooperation-based regional innovation projects (2021RIS-003).

Data Availability Statement: The datasets used and/or analyzed during the current study available from the corresponding author on reasonable request.

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

Author's contributions: All authors are equally contributed in the research work.

References

[1] Smarandache, F. (1999). A unifying field in Logics: Neutrosophic Logic. In *Philosophy* (pp. 1-141). American Research Press.

[2] Kandasamy, I., Vasantha, W. B., Obbineni, J. M., & Smarandache, F. (2020). Sentiment analysis of tweets using refined neutrosophic sets. *Computers in Industry*, *115*, 103180.

[3] Abdel-Basset, M., Chang, V., Mohamed, M., & Smarandache, F. (2019). A refined approach for forecasting based on neutrosophic time series. *Symmetry*, *11*(4), 457.

[4] Guan, H., Dai, Z., Guan, S., & Zhao, A. (2019). A neutrosophic forecasting model for time series based on first-order state and information entropy of high-order fluctuation. *Entropy*, *21*(5), 455.

[5] Singh, P. (2020). A novel hybrid time series forecasting model based on neutrosophic-PSO approach. *International Journal of Machine Learning and Cybernetics*, *11*(8), 1643-1658.

[6] Guan, H., He, J., Guan, S., & Zhao, A. (2019). Neutrosophic soft sets forecasting model for multi-attribute time series. *IEEE Access*, *7*, 25575-25588.

[7] Singh, P. (2020). A novel hybrid time series forecasting model based on neutrosophic-PSO approach. *International Journal of Machine Learning and Cybernetics*, *11*(8), 1643-1658.

[8] Edalatpanah, S. A., Hassani, F. S., Smarandache, F., Sorourkhah, A., Pamucar, D., & Cui, B. (2024). A hybrid time series forecasting method based on neutrosophic logic with applications in financial issues. *Engineering applications of artificial intelligence*, *129*, 107531.

[9] Xu, Y., & Cohen, S. B. (2018, July). Stock movement prediction from tweets and historical prices. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 1970-1979).

[10] Smarandache, F. (1999). A unifying field in Logics: Neutrosophic Logic. In *Philosophy* (pp. 1-141). American Research Press.

[11] Dias, B. C. D., Sadaei, H. J., e Silva, P. C. D. L., & Guimarães, F. G. (2021, May). Aggregation of Sentiment Analysis Index with Hesitant Fuzzy Sets for Financial Time Series Forecasting. In *2021 IEEE World AI IoT Congress (AIIoT)* (pp. 0433-0439). IEEE.

DOI: https://doi.org/10.54216/IJNS.250215
Received: February 14, 2024 Revised: May 02, 2024 Accepted: August 08, 2024