Вісник Тернопільського національного технічного університету
https://doi.org/10.33108/visnyk_tntu
Scientific Journal of the Ternopil National Technical University
2025, № 3 (119) https://doi.org/10.33108/visnyk_tntu2025.03
ISSN 2522-4433. Web: ***visnyk.tntu.edu.ua***

UDC 004

COMPARATIVE ANALYSIS OF MACHINE LEARNING ALGORITHMS FOR MARKET CAPITALIZATION TIME SERIES FORECASTING

Oleksandr Zadvornyi; Oleh Pastukh

Ternopil Ivan Puluj National Technical University, Ternopil, Ukraine

Abstract. *This paper presents a comparative analysis of four time series forecasting algorithms applied to market capitalization data of the world's leading companies: Time Series Transformer, Recurrent Neural Network (RNN), Autoregressive Integrated Moving Average (ARIMA) and Simple Moving Average (SMA). The study is based on monthly market cap data for 1000 companies from 2000 to 2025, collected via Yahoo Finance. In addition to temporal dynamics, static categorical features such as sector, industry and market cap category were considered. The models were evaluated using MASE, sMAPE and MAPE metrics. Results show that the transformer-based model achieved the highest accuracy (MASE = 2.01, sMAPE = 15.63%, MAPE = 17.44%), confirming its suitability for long-term forecasting, especially when categorical features are incorporated. ARIMA and RNN showed moderate performance, while SMA performed the worst. Visualization further confirmed the transformer's ability to capture seasonal patterns and trends. Future work includes integrating macroeconomic indicators to enhance prediction accuracy.*
Key words: *market capitalization, time series forecasting, transformers, ARIMA, SMA, RNN, machine learning, financial data, TimeSeriesTransformer.*

https://doi.org/10.33108/visnyk_tntu2025.03.026 *Received 01.08.2025*

1. INTRODUCTION

In modern financial engineering, particularly in the field of trading, there is an increasing interest in the use of intelligent algorithmic agents to support decision-making related to the buying or selling of financial assets. These agents form decision support systems that assist traders in analyzing market trends and forecasting future values of financial indicators such as company market capitalization [1, 2]. At the core of these agents lie various algorithmic paradigms, including Simple Moving Average (SMA), Autoregressive Integrated Moving Average (ARIMA), Recurrent Neural Networks (RNN) and modern transformers, which have gained popularity due to their effectiveness in processing sequential data [3, 4].

The quality of trading decisions largely depends on the accuracy of forecasts generated by these algorithms. Market capitalization, as a key measure of a company's value, is particularly difficult to predict due to its dependency on economic, market and external factors [5]. Meanwhile, the availability of financial data from open sources such as Yahoo Finance or Bloomberg provides opportunities for building and testing forecasting models [6].

Classical approaches to financial time series forecasting commonly employ ARIMA models, which perform well on stationary processes, but face limitations when dealing with non-stationary or highly volatile series [7]. On the other hand, neural network-based approaches offer greater flexibility and the ability to capture long-term dependencies, though they often require substantial computational resources and time for tuning [8].

Recently, transformers have emerged as state-of-the-art architectures for time series forecasting. Thanks to the attention mechanism, transformers can efficiently model long sequences, parallelize computations and capture contextual relationships between sequence elements regardless of their temporal distance. Furthermore, they can incorporate both temporal and static features, which is especially beneficial for financial forecasting [9]. For instance, the works of Liyilei Su et al. demonstrate the effectiveness of hybrid transformer architectures in multi-horizon forecasting with categorical features [10].

The increasing volume, variability and interdependencies of time series in the financial domain further complicate forecasting tasks. In response, researchers are exploring deeper transformer architectures and adaptive attention mechanisms designed to handle long sequences [11, 12].

This study aims to compare the forecasting accuracy of four algorithms, including SMA, ARIMA, RNN, TimeSeriesTransformer, on the task of predicting market capitalization for the world's leading companies. The research is based on a dataset comprising monthly market cap values for 1000 companies since 2000, retrieved via Yahoo Finance, and incorporates static features such as sector, industry and market capitalization category.

The scientific novelty of this work lies in evaluating the effectiveness of transformer-based models with categorical features, compared to traditional (SMA, ARIMA) and neural (RNN) approaches, using a large real-world financial dataset. The results can be utilized by investors and financial analysts to develop more accurate forecasting models and support well-grounded investment decisions.

2. LITERATURE REVIEW

The problem of time series forecasting in the financial sector has been studied for over half a century. One of the earliest and most widespread approaches has been statistical modeling, particularly the Autoregressive Integrated Moving Average (ARIMA) model. Proposed in the classical work of Box and Jenkins, this model is well-suited for short-term forecasting of stationary or transformed-to-stationary series [7]. ARIMA takes autocorrelations into account but has several limitations, including the need for pre-processing, limited capacity to handle seasonality and low adaptability to structural breaks.

Another simple yet common method is the Simple Moving Average (SMA), which relies on averaging the most recent observations over a fixed window. This approach is widely used in trading strategies due to its simplicity and interpretability. However, SMA does not account for trend components and has limited predictive power in complex, nonlinear tasks such as market capitalization forecasting, where external factors have significant influence [2].

In response to the limitations of classical statistical models, neural network-based methods have gained popularity. Among the most successful architectures for sequence modeling are Recurrent Neural Networks (RNN), especially the Long Short-Term Memory (LSTM) variant [8]. RNNs are capable of retaining information about past states, which is particularly important for modeling long-term dependencies. However, they suffer from vanishing or exploding gradients when trained on long sequences and often require large volumes of data [12].

In recent years, transformer-based architectures have become dominant in time series forecasting. First introduced in the work by Vaswani et al., "Attention is All You Need" (2017), transformers rely on a self-attention mechanism that enables the model to capture relationships between any elements in a sequence regardless of their distance [3]. While initially developed for natural language processing tasks such as translation and text generation, transformers have since been adapted for time series forecasting, including models like Time Series Transformer, Temporal Fusion Transformer and Informer [4, 10].

Research findings demonstrate that transformer models trained on large financial datasets can achieve forecasting accuracy comparable to or even exceeding that of specialized models. A major advantage of transformers lies in their ability to incorporate both static and dynamic features, as well as to scale across a large number of time series. Recent implementations include support for multivariate inputs, masking, positional encoding and regularization techniques to enhance generalization [9, 11].

Particular attention should be given to the systematic literature review by Zhang and Wang [11], in which transformers are categorized by architecture, forecasting functions and training strategies. The study also highlights the need for specialized mechanisms to handle long-term dependencies, volatility in financial time series and the integration of temporal and structural features.

In summary, the current literature emphasizes that transformers are a promising direction for financial time series forecasting. However, their practical deployment requires a deep understanding of the architecture, careful hyperparameter tuning and proper data preparation. At the same time, a comparative analysis with traditional methods such as ARIMA and SMA, as well as with RNNs, enables an objective assessment of the applicability of transformers in specific real-world tasks. Therefore, this study aims to implement all of the aforementioned approaches and conduct an empirical comparison of their forecasting performance using a real-world dataset of market capitalization for leading global companies.

3. MATERIALS AND METHODS

To conduct this study, a complete data pipeline was developed for the collection, preprocessing and preparation of financial time series data from open sources. The data source was Yahoo Finance, accessed via the «yfinance» library [6]. Monthly historical market capitalization data for the top 1000 global companies were collected for the period from January 2000 to May 2025 (Fig. 1). The main selection criteria were the availability of at least 60 months of historical records and the presence of share count data required for market capitalization calculation.

In addition to the time-varying target variable (market capitalization in billions of USD), the dataset incorporated three categorical static features: sector (e.g., Technology, Financials), industry (e.g., Semiconductors, Banks) and market cap category (Small Cap, Mid Cap, Large Cap, Mega Cap), which were defined based on the current market value of each company. The data were cleaned from missing values, encoded and transformed into a unified format compatible with the «GluonTS» and «Hugging Face Datasets» libraries [4, 10]. For each company, the data were split into training, validation and test subsets, where the last 12 months were used for testing and the preceding 12 months for validation. The forecasting horizon for all models was fixed at 12 months.

Figure 1. Example of market capitalization time series for the world's leading companies

Four models were implemented and evaluated for comparative forecasting analysis. The first was the Simple Moving Average (SMA) – a baseline model that computes the forecast as the mean of the last six observations. It does not require training and serves as a reference point. The second was the Autoregressive Integrated Moving Average (ARIMA) – a classical statistical model [7]. Its parameters were selected automatically, considering 12-month seasonality and stationarity was tested using the Augmented Dickey-Fuller (ADF) test. This model was trained independently for each company.

The third model was a Recurrent Neural Network (RNN), specifically a Long Short-Term Memory (LSTM) network implemented in PyTorch. The network consisted of two layers with 64 units each, used the Adam optimizer, and minimized the Mean Squared Error (MSE) loss function. The model was trained on 24-month sliding windows.

Finally, a Time Series Transformer model was implemented – a modern architecture based on attention mechanism (Fig. 2). It uses TimeSeriesTransformerForPrediction class from Hugging Face Transformers [10], with 4 encoder and 4 decoder layers and a model dimension of 32. Importantly, this model was designed to incorporate categorical static features via embedding layers, allowing it to adapt to the specific context of each company.

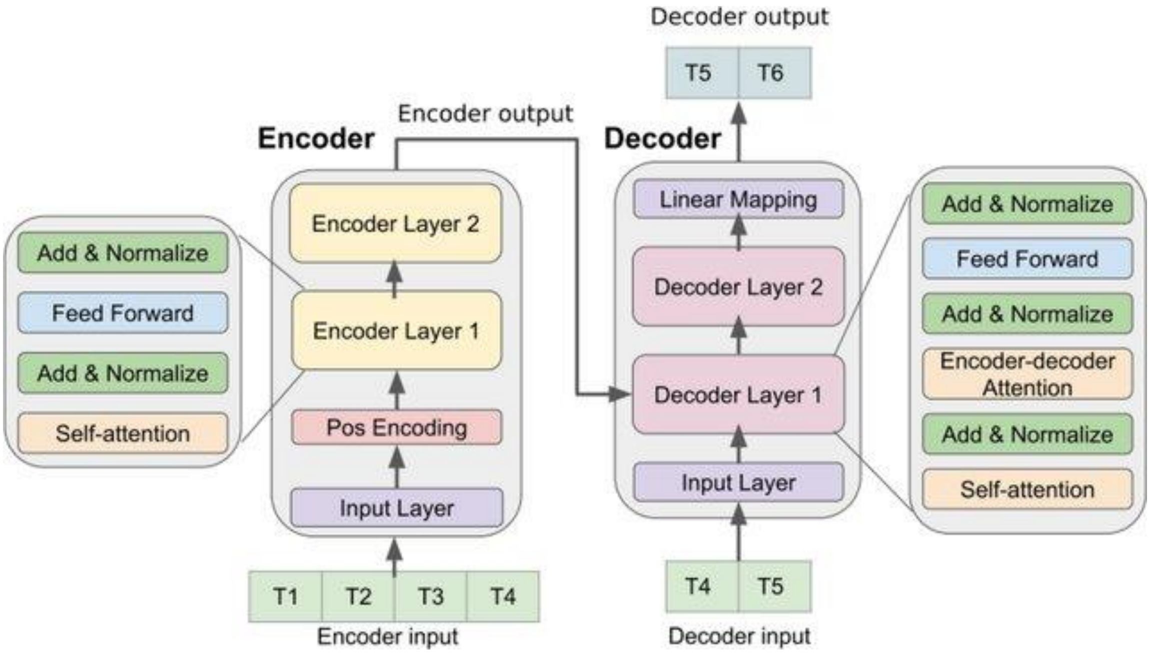

Figure 2. Architecture of the Time Series Transformer model with categorical features [13]

To assess forecasting accuracy, three standard metrics were used. The first was MASE (Mean Absolute Scaled Error), which compares model performance against a naive forecast by normalizing the error. The second was sMAPE (Symmetric Mean Absolute Percentage Error), a symmetric relative error less sensitive to scale. Lastly, MAPE (Mean Absolute Percentage Error) provided a percentage-based interpretation of the error. The use of MASE, sMAPE and MAPE is also recommended in recent studies on financial forecasting and asset allocation evaluation [14].

The experimental setup ensured that all models were trained on the same data splits and evaluated on an independent test set. Forecast plots were also generated for selected companies to aid interpretation (Fig. 3). The entire experimental procedure, from data preparation to plotting and metric calculation, was automated using Python scripts.

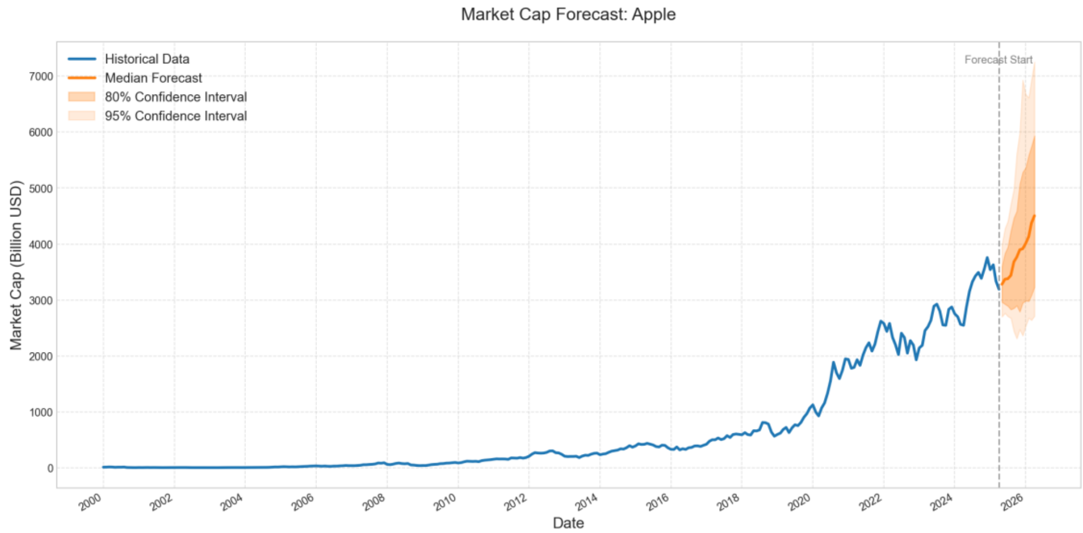

Figure 3. Example forecast for Apple with confidence intervals

This approach enabled objective and reproducible comparison of time series forecasting methods and helped assess the potential of modern architectures in the financial domain. The following section presents the quantitative results of the experiment and the visualized model forecasts.

4. RESULTS

This section presents a comparative analysis of the forecasting performance of four models: SMA, ARIMA, RNN and Time Series Transformer. Evaluation was performed on a 12-month test set that was separated for each company. To ensure comprehensive analysis, the results are reported both as average metric values across all time series and as visualizations of forecasts for selected companies. The table below (Table 1) shows the mean values of MASE, sMAPE and MAPE for each model on the complete test set.

Table 1

Forecasting accuracy comparison of the models.

Model	MASE	sMAPE(%)	MAPE(%)
TimeSeriesTransformer	2.01	15.63	17.44
RNN	2.84	18.22	19.89
ARIMA	3.97	21.45	19.24
SMA	4.02	21.73	19.42

As observed in the table, the transformer-based model achieved the best results, delivering the lowest values across all three metrics. This indicates the model's ability to effectively capture both trend and seasonal components. In contrast, the SMA model performed the worst, highlighting its limited capacity to forecast complex financial patterns.

For a more detailed evaluation, the following figures (Fig. 4–7) illustrate the predicted market capitalization values for selected companies: Apple, Amazon, Tencent and Alibaba. Each plot shows the actual market capitalization over the last 24 months (including the 12-month test period), along with 12-month forecasts generated by all four models. Similar model comparison approaches have been applied in prior studies and in transformer performance assessments for forecasting tasks [15].

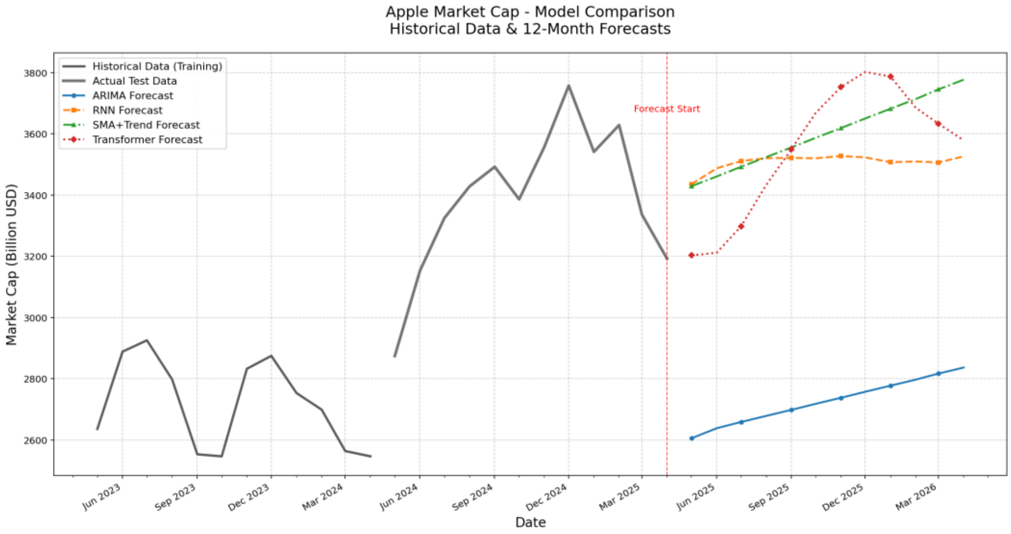

Figure 4. Forecast of Apple's market capitalization (billion USD)

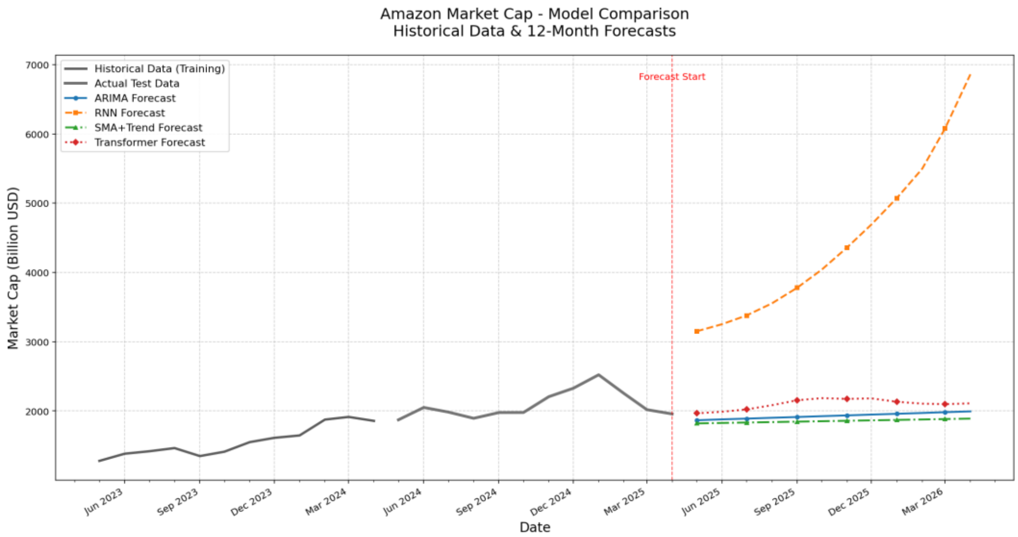

Figure 5. Forecast of Amazon's market capitalization (billion USD)

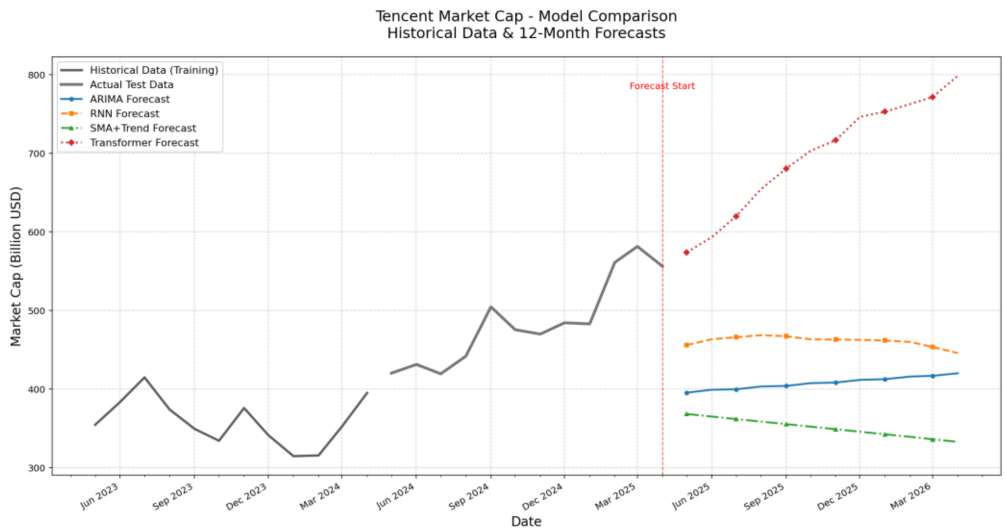

Figure 6. Forecast of Tencent's market capitalization (billion USD)

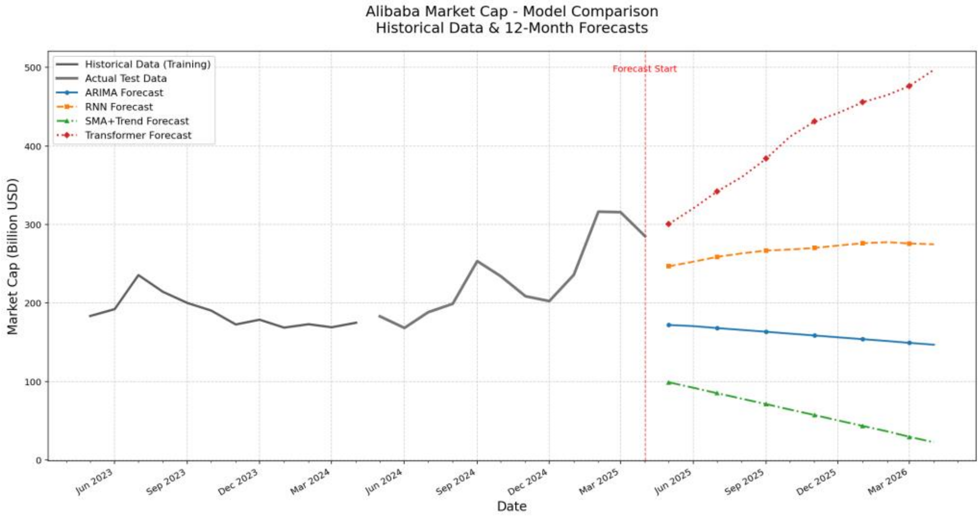

Figure 7. Forecast of Alibaba's market capitalization (billion USD)

The results confirm that the transformer-based model is the most versatile for forecasting market capitalization time series, regardless of industry sector or company size. Additionally, it was found that incorporating categorical features significantly improves forecast accuracy compared to models relying solely on temporal inputs.

5. CONCLUSIONS

This study conducted a comparative analysis of four approaches to forecasting time series of market capitalization for the world's largest companies: Simple Moving Average (SMA), ARIMA, Recurrent Neural Network (RNN) and Transformer (Time Series Transformer). The models were trained and evaluated on a large real-world financial dataset obtained from Yahoo Finance, which included both temporal market capitalization values and static company features: industry, sector and market cap category.

It was found that the transformer-based model outperformed all other tested approaches. Specifically, it achieved the lowest values of MASE, sMAPE and MAPE on the test set, indicating its ability to model both seasonal and trend components of time series. The results show that incorporating categorical (static) features has a positive impact on prediction accuracy, particularly for transformer-based models.

Compared to the transformer, the RNN model demonstrated weaker performance, which can be attributed to its limited capability in modeling long-term dependencies and its less effective use of static features. While ARIMA remains a useful tool for baseline forecasting, it lacks adaptability to structural shifts in time series. The SMA method showed the lowest accuracy, confirming its limited suitability for complex financial scenarios.

These findings suggest that modern attention-based models, particularly transformers, are a promising direction for the development of intelligent financial forecasting systems capable of incorporating both temporal dynamics and contextual characteristics of the analyzed entities. Such systems can generate more accurate and adaptive forecasts, which may benefit investors, financial analysts and algorithmic trading platforms.

Several promising directions are identified for future research. First, the feature set could be extended to include macroeconomic indicators such as GDP, interest rates and inflation. Second, incorporating textual data from news or company reports could enable multimodal forecasting approaches. It would also be valuable to compare the existing models with other modern architectures such as N-BEATS, Temporal Fusion Transformer and DeepAR. Moreover, optimizing model hyperparameters using automated search strategies like

Optuna is recommended. Finally, adapting models for streaming data and real-time forecasting presents another important avenue for further exploration.

In conclusion, the results confirm the high effectiveness of transformer-based models in financial forecasting and highlight the wide potential for advancing hybrid and integrated approaches to modeling financial time series data.

References

1. Ian Goodfellow, Yoshua Bengio, Aaron Courville (2016). Deep Learning, MIT Press, 800 p. ISBN 9780262035613.
2. Yue Chen, Mohan Li (2024) Economic Forces in Stock Returns, arXiv preprint arXiv:2401.04132. https://arxiv.org/abs/2401.04132.
3. Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, Illia Polosukhin (2017,) Attention is all you need, Advances in Neural Information Processing Systems, vol. 30, pp. 5998–6008.
4. Bryan Lim, Sercan Ö. Arık, Nicolas Loeff, Tomas Pfister (2021) Temporal Fusion Transformers for interpretable multi-horizon time series forecasting, International Journal of Forecasting, vol. 37, no. 4, pp. 1748–1764. https://doi.org/10.1016/j.ijforecast.2021.03.012
5. Angelo Ranaldo, Paul Söderlind (2010) Safe haven currencies, Review of Finance, vol. 14, no. 3, pp. 385–407. https://doi.org/10.1093/rof/rfq007
6. Ran Aroussi, yfinance: Python library for Yahoo Finance API, GitHub Repository, 2023. https://github.com/ranaroussi/yfinance.
7. George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, Greta M. Ljung (2015) Time Series Analysis: Forecasting and Control, 5th ed., Wiley, 712 p. ISBN 9781118675021.
8. Benyamin Ghojogh (2023) Recurrent Neural Networks and Long Short-Term Memory Networks: Tutorial and Survey, arXiv preprint arXiv:2304.11461. https://arxiv.org/abs/2304.11461
9. Hajar Gohari, Xuan-Hong Dang, Syed Shah, et al. (2024). Modality-aware Transformer for Financial Time Series Forecasting, Proceedings of the 5th ACM International Conference on AI in Finance, pp. 677–685. https://doi.org/10.1145/3677052.3698654
10. Liyilei Su, Xumin Zuo, Rui Li, et al. (2025) A systematic review for transformer-based long-term series forecasting, Artificial Intelligence Review, vol. 58, article 80. https://doi.org/10.1007/s10462-024-11044-2
11. Zhicheng Zhang, Yong Wang, Shaoqi Tan, et al. (2025) Enhancing Transformer-based models for long sequence time series forecasting via structured matrix, Neurocomputing, vol. 625, pp. 1–12. https://doi.org/10.1016/j.neucom.2025.129429
12. Zhen Fang, Xu Ma, Huifeng Pan, Gonzalo R. Arce (2023) Movement forecasting of financial time series based on adaptive LSTM-BN network, Expert Systems with Applications, vol. 213, 119207. https://doi.org/10.1016/j.eswa.2022.119207
13. Diaa Salman, Cem Direkoglu, Mehmet Kusaf, Murat Fahrioglu (2024) Hybrid deep learning models for time series forecasting of solar power, Neural Computing and Applications, vol. 36, no. 16, https://doi.org/10.1007/s00521-024-09558-5
14. Ying Zhao, Han Sun (2024) Financial asset allocation strategies using statistical and machine learning models, Applied Soft Computing, vol. 150.
15. Vasyl Martsenyuk, Natalia Kit (2024) A multivariate method of forecasting the nonlinear dynamics of production network based on multilayer neural models, Scientific Journal of TNTU, vol. 114, no. 2, pp. 39–50. https://doi.org/10.33108/visnyk_tntu2024.02.039

УДК 004

ПОРІВНЯЛЬНИЙ АНАЛІЗ АЛГОРИТМІВ МАШИННОГО НАВЧАННЯ ДЛЯ ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ РИНКОВОЇ КАПІТАЛІЗАЦІЇ

Олександр Задворний; Олег Пастух

Тернопільський національний технічний університет імені Івана Пулюя, Тернопіль, Україна

Резюме. *Виконано порівняльний аналіз ефективності чотирьох алгоритмів прогнозування часових рядів ринкової капіталізації провідних компаній світу: трансформерів (Time Series Transformer),* *рекурентних нейронних мереж (RNN), авторегресійної інтегрованої ковзної середньої (ARIMA) та простої ковзної середньої (SMA). Дослідження проведено на великому наборі даних, що включає щомісячні значення ринкової капіталізації 1000 компаній з 2000 до 2025 року, зібрані через Yahoo Finance. Враховано не лише часову динаміку, але й статичні ознаки, такі, як сектор, галузь та категорія ринкової капіталізації. Дані опрацьовано та приведено до форматів, сумісних з GluonTS та Hugging Face, з поділом на тренувальну, валідаційну та тестову вибірки. Отримано середні значення трьох метрик точності прогнозу: MASE, sMAPE та MAPE для кожного з алгоритмів. Показано, що модель на основі трансформера демонструє найкращі результати: середнє значення MASE становило 2.01, sMAPE – 15.63%, а MAPE – 17.44%. Звідси випливає, що трансформери є найбільш придатними для задач довготривалого прогнозування, особливо у випадках, коли доступні статичні ознаки, які можна врахувати при навчанні. RNN-модель показала гірші результати, що пояснюється проблемою зникнення градієнта та меншою здатністю до опрацювання довгих залежностей. Модель ARIMA мала помірну точність, проте її ефективність знижувалася на нестаціонарних часових рядах. Найменшу точність прогнозів показала модель SMA, що свідчить про її обмежену придатність для складних фінансових даних. Додатково показано, що включення категоріальних ознак значно покращує продуктивність трансформерів. Візуалізація прогнозів підтвердила, що трансформери здатні надійно моделювати сезонність і тренди в ринковій капіталізації. Перспективним напрямом подальших досліджень є інтеграція макроекономічних факторів і розширення набору ознак для покращення точності прогнозів.*

Ключові слова: *ринкова капіталізація, прогнозування часових рядів, трансформери, ARIMA, SMA, RNN, машинне навчання, фінансові дані, TimeSeriesTransformer.*

https://doi.org/10.33108/visnyk_tntu2025.03.026 *Отримано 01.08.2025*