UDC 004.89
IRSTI 28.23.37

https://doi.org/10.55452/1998-6688-2025-22-4-97-106

**[1]Abdildaeva A.A.,**
PhD, acting Professor, ORCID: 0000-0002-6381-9350,
e-mail: assel.abdildaeva@kaznu.edu.kz
**[1]*Nurtugan G.B.,**
PhD student, ORCID: 0000-0001-8824-9502,
*e-mail: g.nurtugun@gmail.com
**[2]Wojtkiewicz K.**
PhD, assistant Professor, ORCID: 0000-0002-7851-4330,
e-mail: krystian.wojtkiewicz@pwr.edu.pl

[1]Al-Farabi Kazakh National University Almaty, Kazakhstan
[2]Wroclaw University of Science and Technology (WUST) Wroclaw, Poland

# DEVELOPMENT AND OPTIMIZATION OF NEURAL NETWORK MODELS WITH ATTENTION MECHANISMS FOR INTRADAY PRICE FORECASTING FOR EUR/USD

**Abstract**

The study examines the problem of intraday forecasting of the EUR/USD currency pair using various neural network architectures, in particular models integrating attention mechanisms. Three neural network architectures were studied: the basic LSTM model, the LSTM model with the Bahdanau attention mechanism, and the Transformer model with the self-attention mechanism. The experiment was conducted on historical minute data for the period from January 2020 to December 2022. The results showed that attentional models are significantly superior to the basic LSTM architecture. The best results were obtained by the Transformer model (MSE=0.185, MAE=0.297, RMSE=0.431, MAPE=7.3%). A detailed analysis confirmed the stability and accuracy of the Transformer model. The identified advantages of attention models justify their prospects for use in algorithmic trading and require further research to optimize and adapt to real trading conditions. In particular, further research may be aimed at integrating attention models with trading strategies and risk management systems, as well as studying their behavior in the face of sudden changes in market volatility. In addition, it is proposed to explore the possibilities of combining attention architectures with other forecasting methods to increase the overall stability and reliability of forecasts in practical trading.

**Keywords:** intraday forecasting, Forex, exchange rates, EUR/USD, neural networks, attention mechanism, Transformer, LSTM, self-attention.

## Introduction

Accurate forecasting of prices for currency pairs, especially on the scale of intraday trading, is a critically important task for participants in the foreign exchange market. The Forex market, where the EUR/USD currency pair is traded, is characterized by high volatility and significant noise levels, especially over short time intervals such as minutes. This significantly complicates the process of forecasting short-term price movements, setting traders and analysts the task of finding effective tools to capture complex and rapidly changing patterns [1].

Traditional forecasting methods, including statistical approaches (for example, ARIMA or exponential smoothing models), often prove to be insufficiently effective when working with high-frequency Forex data. Even classical neural network models such as LSTM and convolutional networks (CNN), despite their power and ability to detect nonlinear dependencies, are sometimes unable to quickly and accurately adapt to sudden market changes. This is due to their limited ability to focus on the most significant elements of time sequences.

In this regard, this study aims to improve the quality of intraday price forecasting of the EUR/USD currency pair through the use of modern neural network architectures that integrate attention mechanisms. Attention mechanisms that have proven effective in areas such as natural language processing and sequence analysis have the potential to highlight the most significant and relevant segments of historical data. In the context of intraday forecasting, this can lead to a significant increase in prediction accuracy by selectively focusing the model on important time points.

The novelty of the research lies in the fact that the mechanisms of attention, including transformer-type architectures and hybrid models combining CNN and LSTM with attention, are still insufficiently studied in the field of algorithmic trading on the Forex market, especially on small time scales. Most of the research in the field of deep learning applied to finance has traditionally focused on daily and weekly intervals and is mainly focused on the stock market. Intraday Forex data and approaches using attention architectures have been studied much less. Thus, this study is aimed at filling the existing gap, assessing how effectively attention mechanisms can improve the quality of short-term forecasts of the minute interval in the foreign exchange market.

It is assumed that the results of this study will provide new knowledge and clearly demonstrate whether modern neural network architectures with attention mechanisms can significantly exceed the accuracy of traditional approaches based on intraday currency data. This will allow us to better understand the possibilities of applying attention models in the practice of algorithmic trading and lay the foundation for further developments in this area.

Traditional Neural Network Models for Intraday Forex

Long-term short-term memory (LSTM) and similar architectures have become the most popular models for predicting financial time series. LSTM captures time dependencies well and by 2020 has demonstrated superiority over classical methods in the task of short-term price forecasting. Thus, Yıldırım et al., LSTM was used to predict the direction of movement of the EUR/USD pair using a hybrid architecture: one network was trained on macroeconomic indicators, the other on technical indicators [2]. Combining the results of two LSTMs through a special rule made it possible to improve the accuracy of the direction classification in comparison with a single LSTM for all signs. The authors also introduced a neutral «no change» class for small fluctuations and showed an increase in the final profit metric (profit_accuracy) to ~73.6%. This approach reduced the number of erroneous transactions, filtering out insignificant price fluctuations. In another study, a two-input LSTM model was proposed, where fundamental and technical indicators are presented separately; this approach also showed an increase in forecast accuracy compared to models using only one type of feature [3].

Other works have investigated various modifications of recurrent networks. In particular, a comparison of LSTM and its bidirectional version of BiLSTM demonstrated some advantage of Bidirectional-LSTM in the accuracy of forecasting exchange rates [4]. Hossain and colleagues proposed a hybrid GRU-LSTM for predicting the prices of four major pairs (EUR/USD, GBP/USD, etc.) at intraday intervals of 10 and 30 minutes [5]. Their two-layer network contains the GRU layer first, and then the LSTM; this combination surpassed the accuracy of the individual LSTM and GRU models, especially in the 10-minute interval. According to MSE, RMSE and MAE, the hybrid showed the smallest error (for 10-minute forecasts) among all comparative models, and in terms of the coefficient of determination R2 it also turned out to be the best, which indicates its more stable results. These results confirmed the effectiveness of combining different types of RNNs to capture complex patterns of intraday fluctuations. Additionally, it was shown that the complication of the LSTM architecture itself in the form of its multilayer version also increases accuracy: the two-layer model (Stacked LSTM) for the AUD/USD pair surpassed the single-layer LSTM and a number of alternative approaches in terms of error indicators [6].

In addition to recurrent networks, convolutional neural networks (CNNs) were also used to extract local patterns from price ranges. However, the direct application of CNNs or deep multi-layer networks to a financial range can lead to instability due to sensitivity to noise [7]. Therefore, CNN is often combined with RNN: for example, in one approach for Forex, it was proposed to first compress the input data of 1-D CNN, and then process the sequence with LSTM (CNN-LSTM). Such hybrids are designed to combine CNN's ability to recognize short patterns with LSTM's ability to work with long dependencies. For example, Markova et al. (2023) developed a CNN-LSTM auto-encoder model for predicting 5-minute EUR/USD prices, which demonstrated very high forecast accuracy (RMSE≈0.0036 on test data) [8]. In another study, the CNN+LSTM combination reduced the standard error of the forecast by about 9% compared to LSTM alone, while increasing the stability and learning rate of the model (~40% gain in training time) [9]. In addition, approaches are being explored taking into account the macroeconomic context: for example, Pornwattanavichai et al. (2022) proposed the BERTFOREX cascade model, where fundamental indicators (such as inflation, GDP, indices, etc.) are first processed using BERT and an autoencoder and then combined with technical indicators; such a hybrid achieved about 79.4% accuracy in forecasting course directions [10].

It is important to note that the success of the model based on historical data does not always guarantee a practical profit when trading costs are taken into account. A study by Ito et al. analyzed the prediction of intraday returns of currencies using LSTM based on events in the bid stack [11]. The model was good at predicting the direction of change in the mid-exchange rate (mid-quote) in minute increments, ahead of a number of other algorithms. However, when taking into account the spread and commissions, such a predictive model ceased to bring economic benefits, in fact confirming the hypothesis of market efficiency. This indicates a limitation of many purely predictive models: without taking into account market frictions, their results may be overestimated. Nevertheless, the very fact that the neural network is able to detect signals in the order flow is valuable – it indicates the presence of predictable patterns, even if they are not directly monetized in the presence of transaction costs.

It should also be mentioned that some comparative studies question the universality of LSTM. For example, Zafeiriou and Kalles tested several architectures on a short Forex trend and found that a special multi-layer perceptron network simulating a technical analyst worked no worse than LSTM with much less computing time [12]. This highlights that the success of models depends on fine-tuning and structure: increasing complexity does not always automatically lead to a better result without taking into account the specifics of the data.

Attention-oriented and hybrid approaches

Current trends in forecasting financial series are related to the introduction of attention mechanisms, which allow the model to focus on the most informative parts of the sequence. In the context of currency series, attention is able to identify important time steps (for example, reactions to news or certain market phases) and thereby improve the quality of the forecast compared to conventional LSTM/GRU, which «remember» everything. From 2020–2021, works began to appear adapting the successful Transformer architecture from the field of natural language processing to financial forecasting tasks. Transformer is completely based on the self-attention mechanism for modeling dependencies in a sequence and eliminates RNN memory length limitations.

One of the first examples of Transformer's use in intraday Forex is the study by Grądzki and Wójcik [13]. Their goal was to test whether the Transformer architecture could outperform modern deep models based on high-frequency currency data. The authors note that intraday trading in the Forex market using deep learning has not been studied enough, so they first identified the best benchmark (it turned out to be ResNet-LSTM), and then compared Transformer with it. The results showed that Transformer has a high predictive power based on Forex minute data and slightly surpasses the carefully tuned ResNet-LSTM in terms of forecast quality. This is an extremely important conclusion: even taking into account that the improvement was not radical, the fact that the attention model is superior confirms the prospects of this approach. Interestingly, Other authors also note the potential of Transformer models for high-frequency trading: for example, a decoder-free model trained on the EUR/USD and GBP/USD minute series reached a loss function (cross-entropy) value below 0.2, which indicates the possibility of building profitable strategies based on attention architectures [14]. In accordance with these observations, Fischer et al. (2024) note that Transformer is very suitable for currency forecasting, however, in order to be confidently superior to LSTM, it requires the use of multiple input features (i.e., taking into account several simultaneously predicted time series and factors) [15]. Interestingly, in the work of Grądzki and Wójcik, the transformer performed better on larger timeframes (up to 12 hours), whereas there is too much noise in the very small data. In general, this study demonstrates the viability of Transformer for intraday currency forecasting and provides a basis for further improvements (for example, cost accounting in real trading – the authors also conducted a backtest of strategies and took into account commissions when evaluating the model).

Another notable trend is hybrid architectures with attention, combining different types of networks. For example, in the related field of cryptocurrencies, Peng et al. proposed the ACLMC model, which combines a convolutional network and an LSTM with an attention mechanism to predict trends of several cryptocurrencies at once at a high frequency [16]. The key idea was to use multi-frequency data and stabilize the target variable using a triple trend classification. As a result, the built-in attention helped the model extract useful information at different time scales and significantly reduced the number of false trades compared to traditional methods. This hybrid model has improved a number of financial quality metrics and surpassed several basic algorithms without attention. In addition, the use of improved attention mechanisms directly in recurrent layers is justified: for example, the integration of a special GRU-enhanced attention mechanism into the LSTM structure has improved the accuracy and stability of predictions in a volatile stock market [17]. Similar results show that the combination of CNN/LSTM with attention is able to simultaneously take into account short-term patterns and long-term dependencies, which is especially important for noisy intraday series.

Complex Transformer architectures are also emerging-derivatives adapted to time series. Zhao and colleagues conducted a detailed comparison of the classic Transformer with its Informer and Temporal Fusion Transformer (TFT) modifications on several currency pairs (for example, NZD/USD, NZD/CNY, etc.) [18]. The results showed that the most advanced TFT model gave the highest prediction accuracy (coefficient R2 to 0.94) with minimal RMSE/MAE errors. Transformer was inferior in accuracy without improvements, but the alternative Informer model trained and converged much faster, albeit with a slight decrease in accuracy. Thus, there is a trade-off between the quality of the forecast and the computational efficiency of the attention models. In addition, Zhao et al. We found that the inclusion of additional market indicators, such as the VIX volatility index, improves the accuracy of TFT forecasts. This underscores the importance of a multimodal approach: attention mechanisms can easily be expanded to multiple data sources (news, volatility indicators, etc.), which has already been demonstrated in the latest models combining price series with text news signals and opens up opportunities for further improving the quality of intraday forecasts. More complex hybrid schemes based on Transformer have recently been proposed: for example, a dual attention mechanism, where one subnet processes price ranges with improved self-attention (with masking and multi-head), and a parallel subnet analyzes related factors through ConvLSTM+BiGRU with self-attention. Combining such networks has significantly improved forecasting accuracy and surpassed five modern models on a variety of financial datasets [19].

The use of attention layers on top of classic LSTMs is also being explored in the stock market. Thus, a modification of Att-LSTM – LSTM with an optimized attention mechanism was proposed for predicting stock prices, and an improvement in accuracy compared to conventional LSTM was shown [20]. Such results indirectly confirm the value of the attention mechanism for currency ranges as well: focusing on the most significant changes should enhance the model's ability to predict sharp intraday movements in the exchange rate.

Discussion and identified gaps

An analysis of modern literature shows that neural network approaches dominate in forecasting exchange rates, especially on high-frequency data. Classical RNN architectures (LSTM/GRU and their combinations) have been successfully used for short-term and intraday forecasts, surpassing traditional models in accuracy. At the same time, they are not without drawbacks: the limited ability to take into account long-term dependencies and important market features (shocks, news) stimulated the introduction of attention mechanisms and completely new architectures (Transformers). Modern attention models have already shown competitiveness and sometimes superiority over the best recurrent networks, especially in the tasks of classifying the direction of price movement and detecting signals in noisy data. However, the literature on attention methods specifically for intraday Forex is still scarce. Only a few papers directly address this area, whereas most studies either focus on daily intervals or study related markets (stocks, cryptocurrencies). In particular, Grądzki and Wójcik emphasize that deep learning based on intraday Forex data remains an insufficiently researched area. Our review confirms this assessment: although some successes (for example, the use of Transformer) have already been achieved, the mechanisms of attention in the context of intraday currency forecasting are still not fully understood. It is not fully clear how best to integrate attention to account for cross-market effects, news flows, and how these models perform in real trading conditions with slippage and fees.

Thus, there is an obvious gap in research related to the in-depth use of attention mechanisms for high-frequency currency ranges. Filling this gap may include the development of specialized attention architectures or hybrids that take into account the specifics of Forex (for example, daily cycles of trading sessions, spikes in activity when macro statistics are released), and comparing them with already proven LSTM/GRU models. The conducted studies substantiate the prospects of such a path: they demonstrate the advantages of existing approaches and at the same time point out the limitations, overcoming which (through the mechanisms of attention) will be the task of further scientific work. This logically leads to the hypothesis of the current study, aimed at using attention mechanisms to improve the accuracy of the intraday forecast - not approving it in advance, but following from the analyzed achievements and gaps in the literature.

**Materials and methods**

In this study, historical intraday data of the EUR/USD currency pair with a minute frequency for the period from January 2020 to December 2022 were used to evaluate the effectiveness of neural network models. The data has been preprocessed using the following steps:

- Data resampling up to a minute interval (to detect gaps);
- Deleting weekend data (Saturdays and Sundays that appeared during resampling);
- Interpolation of missing values using the linear method and subsequent filling using the forward fill method.

The data was divided into three samples: training (up to and including December 2021), validation (from January to September 2022) and test (from October to December 2022). The signs for the models were the opening, maximum, minimum and closing prices. The closing price was chosen as the predicted indicator.

Model Architecture

The following neural network architectures were considered in the study:

- Basic LSTM: The architecture of the model consisted of two recurrent layers of the LSTM type (Long Short-Term Memory) with dimensions of 64 and 32 neurons, respectively. A 0.2 dropout was applied between the layers to prevent overfitting. The network output used a fully connected layer with a single neuron responsible for predicting the closing price.
- LSTM with the attention mechanism: This model used a single LSTM layer with a dimension of 64 neurons and a dropout level of 0.2. Next, the Bahdanau attention mechanism was used, which allowed the model to dynamically distribute the attention weights for each time step. This mechanism was implemented by two fully connected layers (32 neurons with the tanh activation function and 1 neuron for calculating scoring). The final forecast was formed by summing the LSTM outputs weighted by attention weights.
- Transformer (Encoder-only): This architecture was based entirely on the self-attention mechanism. It used a multi-headed attention with 4 heads and a representation dimension (d_model) of 64. LayerNormalization and positional embeddings were also included in the model. After the

self-attention layer, the Feed-Forward Neural Network block was used (128 neurons with ReLU activation function and additional dropout 0.2 regularization). The final forecast was carried out using a global average pooling layer (GlobalAveragePooling1D) and a fully connected layer with one neuron.

All models were trained using the standard error loss function (MSE) and the Adam optimizer. To prevent overfitting, early stopping (EarlyStopping) and maintaining the best model weights (ModelCheckpoint) were used. The quality of the models was assessed using the MSE, MAE, RMSE and MAPE metrics in a test sample.

Thus, the proposed methodology made it possible to conduct a systematic comparison of neural network architectures and identify the contribution of the attention mechanism to improving the accuracy of intraday price forecasting of the EUR/USD currency pair.

**Results and discussion**

Table 1 – Model metrics based on test data

| Model | MSE | MAE | RMSE | MAPE |
|---|---|---|---|---|
| Basic LSTM | 1.773636 | 1.158659 | 1.331779 | 32.238757 |
| LSTM+Attention | 0.616269 | 0.745160 | 0.785028 | 22.013667 |
| Transformer | 0.185482 | 0.297064 | 0.430676 | 7.299603 |

The results of the experiment show pronounced differences in the effectiveness of the considered neural network architectures. The best performance was demonstrated by the Transformer model, which has a self-attention mechanism, which significantly surpassed the other studied models in all evaluation metrics (Table 1). The mean square error (MSE) of the Transformer model was 0.185, which is significantly less than the LSTM+Attention model (MSE=0.616) and the basic LSTM (MSE=1.774). A similar trend is observed for other key metrics: Transformer showed minimal average absolute errors (MAE=0.297) and standard deviation (RMSE=0.431), as well as the smallest average absolute percentage error (MAPE=7.3%).

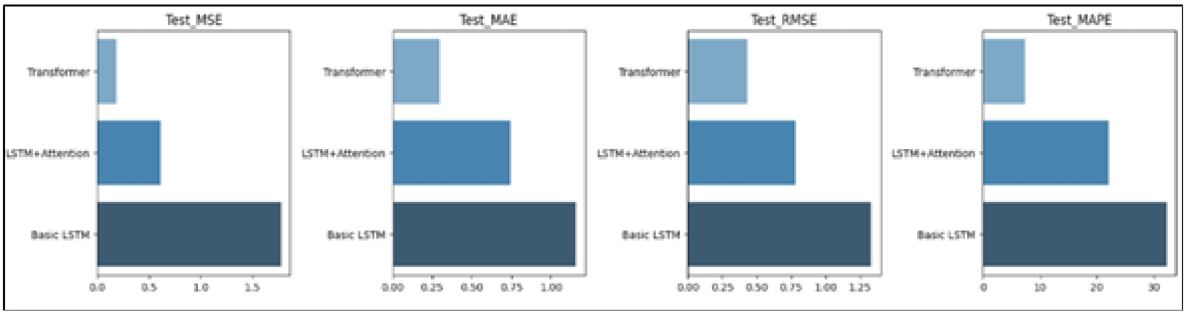


Figure 1 – Comparison of model indicators based on test data

The attentional LSTM model also significantly improved the prediction quality compared to the basic LSTM, demonstrating a significant reduction in errors (MSE=0.616, MAE=0.745, RMSE=0.785, and MAPE=22.01%). The basic LSTM model, on the contrary, performed worse (Figure 1), which indicates its insufficient ability to take into account complex and rapidly changing market conditions.

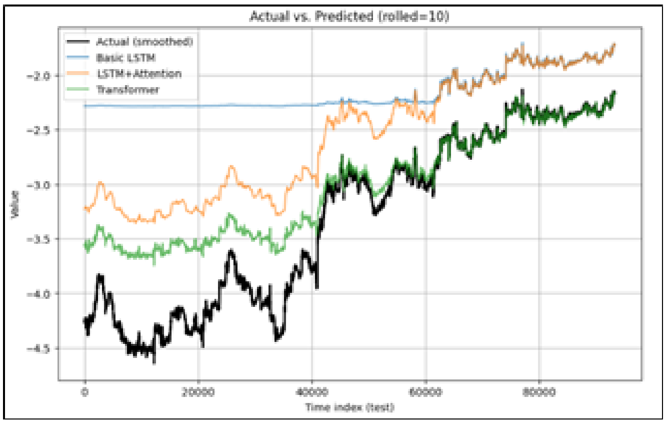


Figure 2 – Model forecasts

An additional analysis of the visual graphs (Figure 2) of the real and predicted values confirmed the numerical results. The predictions of the Transformer model are as close as possible to the real values, while the remnants of its predictions are minimal and do not show a pronounced systematic error (Figure 3). The LSTM model with the attention mechanism also shows stable behavior, but its deviations from real values are noticeably greater compared to Transformer. At the same time, the basic LSTM model systematically overestimates forecasts, as can be seen from both the remainder graph and the error distribution.

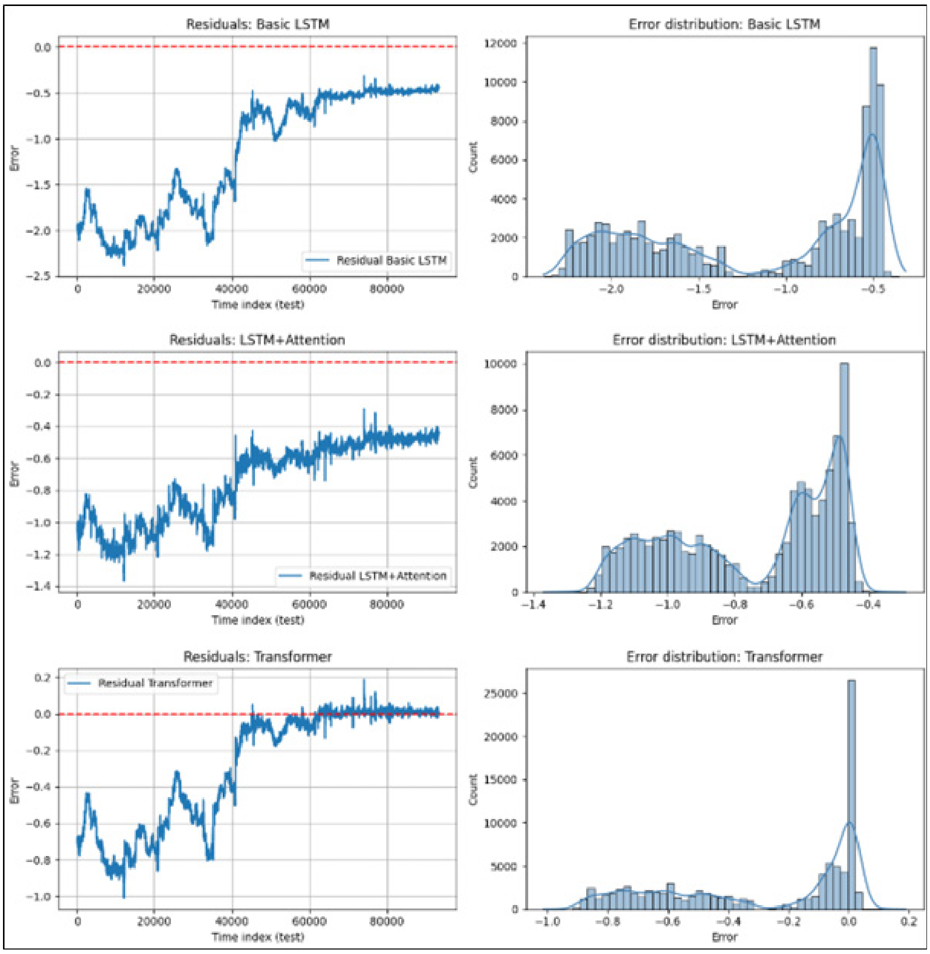


Figure 3 – Error distribution and residuals

The results obtained suggest that the attention mechanisms, in particular, the Transformer architecture, have a high potential for use in the tasks of intraday forecasting of exchange rates. The strong superiority of the Transformer model over the basic LSTM underscores the importance of dynamically managing attention to various segments of historical data and confirms the relevance of adaptive neural network architectures.

It is worth noting that even the LSTM model, complemented by a simple Bahdanau attention mechanism, is significantly superior to the classical recurrent architecture without attention, which confirms the hypothesis of the study on the expediency of using attention mechanisms in forecasting high-frequency financial series.

Nevertheless, the success of the Transformer model is due not only to the attention mechanism itself, but also to the complexity of the architecture, which requires significant computing resources and careful tuning of hyperparameters. This opens up opportunities for further research aimed at optimizing and adapting attention models for practical use in trading. It is also important to take into account that the results may vary depending on currency pairs and time intervals, which requires additional testing and analysis of the resilience of models to changing conditions.

## Conclusion

As a result of the research, it was proved that the use of neural network models with an attention mechanism is a promising direction for the tasks of intraday forecasting of exchange rates. The Transformer model, which implements the self-attention mechanism, demonstrated the best results for all the studied metrics based on the data of the EUR/USD currency pair.

The results obtained confirm the hypothesis that attention mechanisms can significantly improve the accuracy of short-term forecasting based on high-frequency data, opening up prospects for their wide application in algorithmic trading. At the same time, it should be borne in mind that the success of such models depends on their computational efficiency and hyperparameter settings. Future research may be aimed at expanding the number of currency pairs under consideration, exploring various attention architectures and developing methods for their effective configuration.

## REFERENCES

1 Dakalbab, F., Kumar, A., Abu Talib, M., and Nasir, Q. Advancing Forex prediction through multimodal text-driven model and attention mechanisms. Intelligent Systems with Applications, 23, 200518 (2025). https://doi.org/10.1016/j.iswa.2025.200518

2 Yildirim, D.C., Toroslu, I.H., and Fiore, U. Forecasting directional movement of Forex data using LSTM with technical and macroeconomic indicators. Financial Innovation, 7 (1), 1–21 (2021). https://doi.org/10.1186/s40854-020-00220-2

3 Saghafi, A., Bagherian, M., and Shokoohi, F. Forecasting Forex EUR/USD closing prices using a dual-input deep learning model with technical and fundamental indicators. Mathematics, 13 (9), 1472 (2025).

4 García, F., Martínez-César, S., Pérez-Rosés, H., et al. Foreign exchange forecasting models: LSTM and BiLSTM comparison. Engineering Proceedings, 68 (1), 19 (2024).

5 Islam, M.S., and Hossain, E. Foreign exchange currency rate prediction using a GRU–LSTM hybrid network. Soft Computing Letters, 3, 100009 (2021).

6 Ayitey Junior, M., Appiahene, P., and Appiah, O. Forex market forecasting with two-layer stacked long short-term memory neural network (LSTM) and correlation analysis. Journal of Electrical Systems and Information Technology, 9 (1), 14 (2022).

7 Bormpotsis, C., Sedky, M., and Patel, A. Predicting Forex currency fluctuations using a novel bio-inspired modular neural network. Big Data and Cognitive Computing, 7 (3), 152 (2023).

8 Markova, M. Forex time series forecasting using hybrid convolutional neural network/long short-term memory network model. In: International Conference on New Trends in the Applications of Differential Equations in Sciences. Cham: Springer International Publishing, 2022, pp. 295–305.

9 Shi, H., Li, Y., Wang, Z., et al. A CNN–LSTM based deep learning model with high accuracy and robustness for carbon price forecasting: A case of Shenzhen's carbon market in China. Journal of Environmental Management, 352, 120131 (2024).

10 Pornwattanavichai, A., Maneeroj, S., and Boonsiri, S. BERTFOREX: Cascading model for Forex market forecasting using fundamental and technical indicator data based on BERT. IEEE Access, 10, 23425–23437 (2022).

11 Ito, K., Iima, H., and Kitamura, Y. LSTM forecasting foreign exchange rates using limit order book. Finance Research Letters, 47, 102517 (2022).

12 Zafeiriou, T., and Kalles, D. Comparative analysis of neural network architectures for short-term FOREX forecasting. Intelligent Decision Technologies, 18 (3), 2177–2190 (2024).

13 Gradzki, P., and Wojcik, P. Is attention all you need for intraday Forex trading? Expert Systems, 41 (2), e13317 (2024).

14 Kantoutsis, K.T., Mavrogianni, A.N., and Theodorakatos, N.P. Transformers in high-frequency trading. Journal of Physics: Conference Series, 2701 (1), 012134 (2024).

15 Fischer, T., Sterling, M., and Lessmann, S. FX-spot predictions with state-of-the-art transformer and time embeddings. Expert Systems with Applications, 249, 123538 (2024).

16 Peng, P., Liu, X., Zhang, Y., et al. Attention-based CNN–LSTM for high-frequency multiple cryptocurrency trend prediction. Expert Systems with Applications, 237, 121520 (2024).

17 Zhang, J., and Yan, L. GRU-enhanced attention mechanism for LSTM in hybrid CNN–LSTM models for stock prediction. Journal of Global Trends in Social Science, 2 (3), 10–17 (2025). https://doi.org/10.70731/rzvs8j53

18 Zhao, L., and Yan, W.Q. Prediction of currency exchange rate based on transformers. Journal of Risk and Financial Management, 17 (8), 332 (2024).

19 Hadizadeh, A., Tarokh, M.J., and Ghazani, M.M. A novel transformer-based dual attention architecture for the prediction of financial time series. Journal of King Saud University – Computer and Information Sciences, 37, 72 (2025). https://doi.org/10.1007/s44443-025-00045-y

20 Liu, Q., Hu, Y., and Liu, H. Enhanced stock price prediction with optimized ensemble modeling using multi-source heterogeneous data: Integrating LSTM attention mechanism and multidimensional gray model. Journal of Industrial Information Integration, 42, 100711 (2024).

**[1]Абдилдаева А.А.,**
PhD, профессор м.а., ORCID: 0000-0002-6381-9350,
e-mail: assel.abdildaeva@kaznu.edu.kz
**[1]*Нуртуган Г.Б.,**
докторант, ORCID: 0000-0001-8824-9502,
*e-mail: g.nurtugun@gmail.com
**[2]Войткевич К.,**
PhD, доцент, ORCID: 0000-0002-7851-4330,
e-mail: krystian.wojtkiewicz@pwr.edu.pl

[1]Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы қ., Қазақстан
[2]Вроцлав ғылым және технологиялар университеті (WUST), Вроцлав қ., Польша

## EUR/USD ЖҰБЫ БОЙЫНША ТӘУЛІКТІК БАҒА БОЛЖАУҒА АРНАЛҒАН НАЗАР МЕХАНИЗМДЕРІ БАР НЕЙРОНДЫҚ ЖЕЛІ МОДЕЛЬДЕРІН ДАМЫТУ ЖӘНЕ ОҢТАЙЛАНДЫРУ

**Аңдатпа**

Зерттеу әртүрлі нейрондық желілік архитектураларды, атап айтқанда назар аудару механизмдерін біріктіретін модельдерді (attention) пайдалана отырып, EUR/USD валюта жұбының бағамын бір күндік болжау мәселесін қарастырады. Үш нейрондық желілік архитектура зерттелді: LSTM базалық моделі, bahdanau зейін механизмі бар LSTM моделі және өзін-өзі зейіндеу механизмі бар трансформатор моделі (self-attention). Эксперимент 2020 жылдың қаңтарынан 2022 жылдың желтоқсанына дейінгі кезеңдегі

тарихи минуттық деректерде жүргізілді. Нәтижелер назар аудару механизмі бар модельдер LSTM негізгі архитектурасынан айтарлықтай жоғары екенін көрсетті. Ең жақсы нәтижелер Transformer моделімен алынды (MSE=0.185, MAE=0.297, RMSE=0.431, MAPE=7.3%). Егжей-тегжейлі талдау transformer моделінің тұрақтылығы мен дәлдігін растады. Attention модельдерінің анықталған артықшылықтары олардың алгоритмдік саудада қолдану перспективасын негіздейді және нақты сауда жағдайларын оңтайландыру және бейімдеу үшін қосымша зерттеулерді қажет етеді. Атап айтқанда, қосымша зерттеулер attention модельдерін сауда стратегиялары мен тәуекелдерді басқару жүйелерімен біріктіруге, сондай-ақ нарықтық құбылмалылықтың күрт өзгеруі жағдайында олардың мінез-құлқын зерттеуге бағытталуы мүмкін. Сонымен қатар, практикалық саудадағы болжамдардың жалпы тұрақтылығы мен сенімділігін арттыру үшін attention архитектураларын басқа болжау әдістерімен біріктіру мүмкіндіктерін зерттеу ұсынылады.

**Тірек сөздер:** тәуліктік болжау, Forex, айырбас бағамдары, EUR/USD, нейрондық желілер, назар механизмі, Transformer, LSTM, өзіндік назар (self-attention).

**[1]Абдилдаева А.А.,**
PhD, и.о. профессора, ORCID: 0000-0002-6381-9350,
e-mail: assel.abdildaeva@kaznu.edu.kz
**[1]*Нуртуган Г.Б.,**
докторант, ORCID: 0000-0001-8824-9502,
*e-mail: g.nurtugun@gmail.com
**[2]Войткевич К.,**
PhD, доцент, ORCID: 0000-0002-7851-4330,
e-mail: krystian.wojtkiewicz@pwr.edu.pl

[1]Казахский национальный университет им. аль-Фараби, г. Алматы, Казахстан
[2]Вроцлавский университет науки и технологий (WUST), г. Вроцлав, Польша

## РАЗРАБОТКА И ОПТИМИЗАЦИЯ НЕЙРОННЫХ СЕТЕВЫХ МОДЕЛЕЙ С МЕХАНИЗМАМИ ВНИМАНИЯ ДЛЯ ВНУТРИДНЕВНОГО ПРОГНОЗИРОВАНИЯ ЦЕНЫ ПАРЫ EUR/USD

**Аннотация**

В исследовании рассматривается проблема внутридневного прогнозирования курса валютной пары EUR/USD с использованием различных нейросетевых архитектур, в частности моделей, интегрирующих механизмы внимания (attention). Были исследованы три нейросетевые архитектуры: базовая модель LSTM, модель LSTM с механизмом внимания Bahdanau и модель Transformer с механизмом самовнимания (self-attention). Эксперимент проводился на исторических минутных данных за период с января 2020 по декабрь 2022 гг. Результаты показали, что модели с механизмом внимания значительно превосходят базовую архитектуру LSTM. Наилучшие результаты были получены моделью Transformer (MSE=0.185, MAE=0.297, RMSE=0.431, MAPE=7.3%). Подробный анализ подтвердил стабильность и точность модели Transformer. Выявленные преимущества attention-моделей обосновывают их перспективность для применения в алгоритмической торговле и требуют дальнейших исследований для оптимизации и адаптации к реальным торговым условиям. В частности, дальнейшие исследования могут быть направлены на интеграцию attention-моделей с торговыми стратегиями и системами управления рисками, а также изучение их поведения в условиях резких изменений рыночной волатильности. Кроме того, предлагается исследовать возможности комбинирования attention-архитектур с другими методами прогнозирования для повышения общей устойчивости и надежности прогнозов в практическом трейдинге.

**Ключевые слова:** внутридневное прогнозирование, Forex, валютные курсы, EUR/USD, нейронные сети, механизм внимания, Transformer, LSTM, самовнимание.

Article submission date: 30.06.2025