Deep learning architectures for financial time-series forecasting. A comparative methodological review

M. CZYŻEWSKA
ORCID: 0009-0009-7829-8779

maja.czyzewska@wat.edu.pl

Military University of Technology, Faculty of Cybernetics,
Kaliskiego 2 Str., 00-908 Warsaw, Poland

This paper presents a comparative review of selected deep learning architectures used in financial time-series forecasting, with particular emphasis on Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNNs), and attention-based hybrid models. The discussion is framed around three representative market environments: cryptocurrency, equity index, and foreign exchange data, in order to highlight how differences in volatility, liquidity, and market structure may affect model suitability. The paper outlines the main challenges of financial forecasting, including nonlinear dynamics, temporal heterogeneity, and sensitivity to changing market conditions. Particular attention is given to hybrid and attention-based approaches, especially the Regularized Self-Attention Regression (RSAR) model, as an example of a more adaptive forecasting framework. The review suggests that no single architecture should be regarded as universally optimal for all financial series. Instead, the suitability of a given model depends on the statistical properties of the analyzed data, the forecasting horizon, and practical constraints such as computational complexity and sensitivity to hyperparameter selection.

Keywords: financial forecasting, deep learning, hybrid models.

DOI: 10.5604/01.3001.0055.7792

1. Introduction

Financial markets are characterized by high volatility, nonlinear dynamics, and frequent abrupt changes associated with macroeconomic announcements and geopolitical events [22], [23], [24]. Under such conditions, accurate forecasting of asset prices remains an important task for both large financial institutions managing diversified portfolios and individual market participants increasingly relying on automated trading strategies. Traditional econometric models, such as ARIMA, still constitute useful analytical tools, yet their ability to capture complex long-term dependencies and sudden structural shifts is limited. As a result, the growing complexity of financial time series has encouraged researchers to seek more flexible predictive approaches based on deep learning [1], [2], [16], [20].

In recent years, deep neural networks have become one of the most important methodological directions in financial forecasting. In particular, Long Short-Term Memory (LSTM) networks have gained considerable attention due to their ability to model sequential dependencies and mitigate the vanishing gradient problem that affects classical recurrent architectures [3], [4]. Their usefulness in financial forecasting has also been confirmed in empirical studies on stock market prediction [10]. Convolutional Neural Networks (CNNs), although originally developed for image analysis, have also proven useful in time-series prediction by extracting local temporal patterns and reducing sensitivity to noise through shared filters and pooling operations [5], [6]. Their broader applicability to sequential data and related forecasting problems has also been discussed in the literature [9]. The literature further shows that hybrid CNN-LSTM architectures can provide better predictive performance than single-model approaches by combining local feature extraction with long-term sequence modeling [7].

A further step in the development of forecasting architectures has been the introduction of attention-based mechanisms derived from transformer models. Attention allows the network to dynamically assign importance to particular observations within the input sequence, emphasizing informative patterns while suppressing less relevant ones [8]. This is especially valuable in financial forecasting, where not all past observations contribute equally to future price formation. One representative example is the Regularization Self-Attention Regression (RSAR) model, which combines LSTM, self-attention, and CNN components in a single framework [2]. Such architectures are designed to capture both long-range temporal dependencies and local structures while simultaneously improving the model's ability to focus on the most informative fragments of the sequence.

For this reason, a comparative review of the main deep learning architectures used in financial time-series forecasting is justified [16], [20]. The aim of this paper is to discuss and compare three important groups of models applied in this area: LSTM networks, CNN-based models, and attention-based architectures, including hybrid solutions. Particular emphasis is placed on their theoretical foundations, practical strengths, limitations, and suitability for different forecasting conditions. To provide a consistent comparative context, the discussion refers to three representative financial market environments: foreign exchange, equity index, and cryptocurrency data. These environments are treated here as general forecasting contexts rather than as separate objects of detailed market analysis.

The remainder of this paper is organized as follows. The next section outlines the main challenges of forecasting financial time series. The following sections present LSTM models, CNN-based approaches, and attention-based architectures used in financial forecasting. The paper then develops a comparative methodological discussion and closes with concluding remarks.

2. Challenges of financial time-series forecasting

Financial time-series forecasting is difficult because market prices are influenced by nonlinear dynamics, abrupt changes, and instability over time. In such settings, methods based on relatively simple statistical assumptions may lose effectiveness, especially when the analyzed series contain both complex dependencies and changing market conditions. This partly explains why deep learning models have become increasingly important in financial prediction tasks and why they are often treated as a promising alternative to more traditional approaches [1], [10], [16], [21].

An additional difficulty follows from the well-known stylized facts of financial data. Financial series are often characterized by non-stationarity, structural breaks, heavy-tailed distributions, and volatility clustering, which means that models fitted to one market regime may lose accuracy when the underlying conditions change. For this reason, forecasting methodology should be assessed not only in terms of in-sample fit, but also with respect to robustness and generalization under changing market conditions.

Another challenge follows from the temporal structure of financial data itself. Financial series usually contain both short-term local fluctuations and longer sequential dependencies. These two properties are not captured equally well by all architectures: LSTM models were introduced specifically to improve the learning of long-term dependencies in sequential data [3], [4], whereas convolutional models are mainly useful for extracting local patterns and short-range structures [9]. Hybrid CNN-LSTM architectures emerged precisely because these capabilities are complementary rather than interchangeable [7].

A separate methodological problem concerns model capacity and hyperparameter selection. Increasing the number of layers, units, or filters may improve model flexibility, but it may also make the network more prone to overfitting, especially when the available dataset is limited. This issue is emphasized in studies on financial forecasting, where optimized hyperparameter tuning can substantially improve predictive performance, while poorly chosen parameters may lead to excessive fitting of training data and weaker generalization [11].

For this reason, the evaluation of forecasting models should not be reduced to predictive accuracy alone. In practice, a useful architecture must combine good forecasting performance with robustness to unseen data and changing market conditions. This need has encouraged the development of more advanced hybrid and attention-based models, including RSAR, which attempt to integrate sequential memory, local feature extraction, and mechanisms for emphasizing the most informative observations within the input sequence [2], [7].

3. LSTM models in financial forecasting

Long Short-Term Memory (LSTM) networks are among the most widely used deep learning architectures in financial time-series forecasting. Their importance follows from the sequential nature of financial data, in which the order of observations carries essential information. Unlike feedforward neural networks, which process inputs without preserving temporal context, recurrent architectures are designed to incorporate information from previous time steps. However, classical recurrent neural networks often struggle with learning long-term dependencies due to the vanishing gradient problem. This limitation was identified in early studies on recurrent learning and became one of the main motivations for the development of LSTM networks as a more stable sequential architecture [3], [4].

The core advantage of LSTM lies in its gated memory mechanism. Instead of relying solely on a simple recurrent state, the architecture introduces a cell state together with three control gates: the forget gate, the input gate, and the output gate. In methodological terms, this allows the network to regulate which information should be preserved, which should be updated, and which should be passed forward at a given stage of sequence processing. Such a mechanism is particularly useful in financial forecasting, where historical data often contain both meaningful long-range dependencies and short-lived noisy fluctuations. As a result, LSTM models are commonly treated as a natural choice for problems involving stock prices, exchange rates, commodities, and other time-dependent financial series [3], [10].

From a formal perspective, the LSTM mechanism can be expressed through a sequence of gated operations applied at each time step. Let x_t denote the input vector, h_{t-1} the previous hidden state, and C_{t-1} the previous cell state. The main gated operations of the LSTM cell can be written as follows (1)–(6). The forget gate determines which information from the previous memory should be retained:

$$f_t = \sigma\left(W_f[h_{t-1}, x_t] + b_f\right) \tag{1}$$

The input gate controls which new information should be introduced into the memory state:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{2}$$

At the same time, the candidate cell state is computed as:

$$\tilde{C}_t = \tanh\,(W_c[h_{t-1}, x_t] + b_c) \tag{3}$$

These components are then combined to update the cell state:

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{4}$$

Finally, the output gate determines which information is passed to the hidden representation:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{5}$$

and the hidden state is obtained as:

$$h_t = o_t \odot \tanh\,(C_t) \tag{6}$$

In practical terms, these operations make it possible to maintain a more stable memory of relevant sequential information across time steps. This property distinguishes LSTM from simpler recurrent architectures and helps explain its usefulness in financial forecasting tasks, in which informative signals may be distributed across longer observation windows.

The practical usefulness of LSTM in financial forecasting has been confirmed in several empirical studies. Fischer and Krauss showed that LSTM-based models can outperform more traditional machine learning approaches in stock market prediction, indicating that long-term sequential memory may capture subtle dependencies not easily recognized by models without recurrent structure [10]. Similar conclusions were reported in studies comparing LSTM with classical econometric approaches. For example, research on gold-price prediction demonstrated that LSTM achieved better accuracy than ARIMA, especially when longer-term and nonlinear patterns had to be modeled [1]. Other studies also showed that LSTM can outperform regression-based approaches in stock price prediction when its parameters are selected appropriately [12]. At the same time, the literature indicates that model performance depends strongly on hyperparameter tuning, since poorly calibrated LSTM architectures may become overfitted or fail to generalize well [11], [12].

Despite these strengths, LSTM models are not free from limitations. Their sequential design improves memory, but it also increases architectural complexity and training sensitivity. As a result, LSTM networks typically require careful tuning of the number of units, the number of layers, the learning process, and regularization-related parameters. They may also become less effective when short-term local patterns play a dominant role in the data, since such structures are not always captured as efficiently by a purely recurrent model. For this reason, although LSTM remains one of the key reference architectures in financial forecasting, it should be treated as a strong but context-dependent solution rather than a universally optimal one [7], [11].

4. CNN models in financial forecasting

Convolutional Neural Networks (CNNs) constitute an important class of deep learning architectures that can also be adapted to financial time-series forecasting. Although they were originally developed for image analysis, their underlying mechanism can be transferred to sequential data through one-dimensional convolutions. The historical roots of CNNs go back to early hierarchical pattern-recognition architectures and later developments in image--processing tasks, where convolutional filters proved effective in extracting local structures from raw input [5][6]. When applied to time series, the same principle allows filters to move along the input sequence and transform raw observations into feature maps that emphasize recurring short-range patterns in the data [9][13].

From the methodological perspective, CNNs offer several properties that are attractive in financial forecasting. First, they allow automatic feature extraction without the need for extensive manual engineering of local patterns. Second, parameter sharing reduces the number of trainable coefficients compared with fully connected architectures. Third, pooling operations reduce dimensionality and help retain the most informative responses of the filters. In practical terms, these mechanisms may shorten training time, reduce hardware requirements, and make CNN-based models computationally efficient in comparison with more complex sequential or hybrid solutions. This is particularly relevant in financial applications, where noisy data and limited computational resources may make simpler architectures more attractive [9], [13], [18].

In financial applications, CNNs are especially useful, when predictive information is concentrated in recent observations or in short-term regularities rather than in long-range temporal dependencies. This makes them suitable for detecting abrupt changes, local price formations, and repeated short-window behaviors that may be difficult to model with simpler approaches. Studies on time-series prediction have shown that CNN-based architectures can remain competitive in such settings, especially when the forecasting task depends strongly on short-horizon structure [13], [18]. At the same time, CNNs are less naturally suited to representing long-term sequential context, since a convolutional layer efficiently captures what happens within a local neighborhood of observations, but does not explicitly preserve memory over long horizons in the same way as recurrent architectures [7].

For this reason, CNNs are often combined with recurrent components in hybrid architectures. In such models, convolutional layers first extract informative features from the input sequence, after which recurrent layers process these features across time. Lu et al. reported, that CNN-LSTM models can achieve better predictive accuracy than standalone CNN or LSTM models, suggesting that the two approaches should be treated as complementary rather than mutually exclusive [7]. Thus, although CNN-based models offer clear practical advantages in terms of simplicity and computational efficiency, their usefulness in financial forecasting remains strongly dependent on whether local temporal structure is sufficient for the analyzed prediction task.

5. Attention-based and hybrid models in financial forecasting

Hybrid architectures became important in financial forecasting because sequential market data often contain more than one type of predictive structure. Some dependencies unfold over longer horizons, whereas others appear as short-lived local formations. This observation motivated the development of CNN-LSTM models, in which convolutional layers first transform the raw sequence into a more informative feature representation and recurrent layers then process these features across time. Empirical studies have shown that such architectures can outperform stand-alone CNN and LSTM models in stock-price forecasting tasks, suggesting that local feature extraction and sequential modeling should often be treated as complementary rather than competing mechanisms [7].

However, even hybrid CNN-LSTM models do not explicitly determine which parts of the input sequence should matter more for the final prediction. This limitation is particularly relevant in financial forecasting, where not all past observations carry equal predictive value. Market series are noisy, event-driven, and sensitive to abrupt disruptions, which means that a useful model should not only process the sequence but also identify which time steps are more informative than others. This is the point at which attention-based mechanisms become methodologically important [19].

The attention mechanism was introduced in a general sequence-modeling setting by Vaswani et al. in Attention Is All You Need [8]. Its key idea is to assign adaptive weights to elements of the input sequence so that the model can emphasize the most relevant information and suppress less informative observations [19]. In the financial context, such a mechanism is attractive because it allows the model to respond selectively to particular events, movements, or local patterns within the input window instead of treating all observations as equally important [8], [13].

A representative example of this design logic is the Regularized Self-Attention Regression (RSAR) model proposed by Zhou et al. [2]. RSAR does not rely on attention as a stand-alone solution. Instead it inserts a self-attention layer between LSTM-based temporal encoding and convolutional feature extraction. In this way, the model first constructs a sequential representation, then reweights it according to estimated relevance, and only afterwards extracts local structures from the attention-weighted representation. This makes RSAR more than a simple hybrid: it is an architecture built around selective prioritization inside the sequence [2].

An additional feature that distinguishes RSAR from simpler attention-based designs is explicit regularization. The model uses L1 regularization for bias terms, L2 regularization for weight matrices, and an additional penalty applied directly to the attention map. The purpose of this solution is not only to reduce overfitting, but also to prevent the learned attention distribution from collapsing into an excessively narrow concentration on one dominant time step. In methodological terms, this makes RSAR a more controlled forecasting framework than architectures that merely add attention without regulating its behavior [2].

A further methodological extension of attention-based hybrid models concerns optimization-assisted architecture design. Since models such as RSAR contain multiple interacting hyperparameters, their performance may depend not only on the structure itself, but also on the way the configuration is selected. Recent methodological work has therefore suggested combining attention-based forecasting architectures with systematic hyperparameter optimization, including multi-objective approaches that seek a balance between predictive accuracy and generalization [15]. From a broader perspective, this suggests that the development of financial forecasting models increasingly involves not only new architectural components, but also more advanced strategies for adapting these components to the statistical properties of a given dataset.

6. Comparative methodological discussion

From a methodological perspective, the reviewed architectures differ not only in structure, but also in the practical conditions under which they are likely to be most useful. LSTM networks remain one of the most natural choices for financial time-series forecasting because their sequential design is well aligned with data in which temporal order is essential. Their ability to preserve long-range dependencies makes them particularly attractive in problems where relevant information may be distributed across longer horizons rather than concentrated only in the most recent observations. At the same time, the literature suggests that the practical effectiveness of LSTM models depends strongly on the quality of model configuration and training design. Although this is true for neural networks more generally, it becomes especially important in architectures with greater representational capacity, where poorly chosen hyperparameters may lead to unstable training, overfitting, or weak generalization [10], [11].

CNN-based models represent a different methodological trade-off. Compared with recurrent and hybrid architectures, they are typically less complex and computationally lighter. Because convolutional filters share parameters and pooling operations reduce dimensionality, CNNs may require fewer trainable parameters and lower computational effort than more elaborate sequential models. From a practical point of view, this makes them attractive when shorter training time, lower hardware requirements, or implementation simplicity are important priorities. Their main limitation is that they are more naturally suited to local temporal structure than to long-range sequential dependencies. Nevertheless, when the most informative patterns are concentrated in short windows of observations, CNN-based models may constitute an efficient and methodologically justified choice [7], [9], [13], [18].

Attention-based and hybrid architectures, including models such as RSAR, introduce a more complex and potentially more adaptive forecasting framework. By combining multiple architectural principles, such as sequential encoding, local feature extraction, and selective weighting of observations, these models offer a richer representation of financial time series than simpler baseline approaches.

In methodological terms, this suggests a higher capacity to capture heterogeneous temporal dependencies and to adapt to situations in which predictive relevance is unevenly distributed within the input window. At the same time, however, this increased flexibility is accompanied by greater architectural complexity, stronger dependence on regularization design, and greater sensitivity to hyperparameter selection. As a result, such models may offer substantial predictive potential, but they also tend to require more computational power, longer training time, and more careful configuration than simpler architectures [2], [8], [14], [19].

Taken together, these observations suggest that the choice of forecasting architecture should not be based solely on theoretical predictive power. It should also depend on the statistical characteristics of the analyzed financial instrument, including volatility, noise level, temporal structure, and the expected forecasting horizon, as well as on practical constraints such as available computational resources and acceptable implementation cost. In this sense, model selection in financial forecasting is not only a question of accuracy, but also of feasibility and robustness.

An additional issue concerns the growing difficulty of manual hyperparameter selection in more complex neural forecasting models. Although deep learning architectures have been studied increasingly extensively, their behavior often remains difficult to interpret in a fully transparent way, which limits the usefulness of purely expert-based manual tuning. This is one of the reasons why optimization-assisted strategies, including multi-objective approaches, are becoming increasingly important in forecasting methodology. Such approaches may help reduce the time cost of model calibration and support the search for configurations that better balance predictive accuracy and generalization than trial-and-error procedures [15].

7. Conclusion

This paper presented a comparative review of selected deep learning architectures used in financial time-series forecasting, with particular emphasis on LSTM networks, CNN-based models, and attention-based hybrid approaches. The discussion showed that these architectures differ not only in structure, but also in the type of temporal information they are most naturally suited to represent and in the practical requirements associated with their use.

The review also indicated that no single architecture should be regarded as universally optimal for all financial forecasting tasks. Instead, the suitability of a given model depends on the statistical properties of the analyzed series, the forecasting horizon, and practical constraints such as computational cost, training complexity, and sensitivity to hyperparameter selection. In this sense, model choice in financial forecasting should be treated as a context-dependent methodological decision rather than as a search for one universally superior solution.

The considerations presented in this paper provide a theoretical and methodological basis for a broader comparative research framework. Further work may extend this review perspective through empirical evaluation of selected architectures under a unified experimental setup across different financial market environments.

8. Bibliography

[1] Madhika Y.R., Kusrini K., Hidayat T., "Gold Price Prediction Using ARIMA and LSTM Models", *Sinkron*, Vol. 7, No. 3, 1255–1264 (2023).

[2] Zhou J., He Z., Song Y., Wang H., Yang X., Lian W., "Precious metal price prediction based on deep regularization Self-Attention Regression", *IEEE Access*, Vol. 8, 2178–2187 (2020).

[3] Hochreiter S., Schmidhuber J., "Long short-term memory", *Neural Computation*, Vol. 9, No. 8, 1735–1780 (1997).

[4] Bengio Y., Simard P., Frasconi P., "Learning long-term dependencies with gradient descent is difficult", *IEEE Transactions on Neural Networks*, Vol. 5, No. 2, 157–166 (1994).

[5] Fukushima K., "Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position", *Biological Cybernetics*, Vol. 36, 193–202 (1980).

[6] LeCun Y., Bottou L., Bengio Y., Haffner P., "Gradient-Based Learning Applied to Document Recognition", *Proceedings of the IEEE*, Vol. 86, No. 11, 2278–2324 (1998).

[7] Lu W., Li J., Li Y., Sun A., Wang J., "A CNN-LSTM-Based Model to Forecast Stock Prices", *Complexity*, Vol. 2020, 1–10 (2020).

[8] Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser L., Polosukhin I., "Attention is all you need", *Advances in Neural Information*

Processing Systems, Vol. 30, 5998–6008 (2017).

[9] Li Z., Liu F., Yang W., Peng S., Zhou J., "A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects", *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 33, No. 12, 6999–7019 (2022).

[10] Fischer T., Krauss C., "Deep Learning with Long Short-Term Memory Networks for Financial Market Predictions", *European Journal of Operational Research*, Vol. 270, No. 2, 645–669 (2018).

[11] Gülmez B., "Stock price prediction with optimized deep LSTM network with artificial rabbits optimization algorithm", *Expert Systems with Applications*, Vol. 227, 120346–120346 (2023).

[12] Li J., "Comparison of Regressions and Multi-feature LSTM in Stock Price Prediction for Top Car Companies", *Advances in Economics, Management and Political Sciences*, Vol. 85, 170–180 (2024), DOI: 10.54254/2754-1169/85/20240869.

[13] Velastegui R., Zhinin-Vera L., Piliza G.E., Chang O., "Time Series Prediction by Using Convolutional Neural Networks", *Proceedings of the Future Technologies Conference (FTC) 2020*, p. 499–511, Springer, Cham 2020.

[14] Zhang Q., Qin C., Zhang Y., Bao F., Zhang C., Liu P., "Transformer-based attention network for stock movement prediction", *Expert Systems with Applications*, Vol. 202, 117239–117239 (2022).

[15] Czyżewska M., "Multi-Objective Optimization of Regularized Self-Attention Regression Models Using NSGA-II: A Methodological Framework", *Proceedings of the 46th International Business Information Management Association Computer Science Conference* (*IBIMA*), *Research in Advancements in Generative AI, Quantum Computing and Computer Security*, Khalid S. Soliman (ed.), p. 559– 566, IBIMA, 2025.

[16] Sezer O.B., Gudelek M.U., Ozbayoglu A.M., "Financial time series forecasting with deep learning: A systematic literature review: 2005–2019", *Applied Soft Computing*, Vol. 90, 106181–106181 (2020).

[17] Song X., Deng L., Wang H., Zhang Y., He Y., Cao W., "Deep learning-based time series forecasting", *Artificial Intelligence Review*, Vol. 58, Art. 23 (2025).

[18] Ismail Fawaz H., Forestier G., Weber J., Idoumghar L., Muller P.-A., "Deep learning for Time Series Classification: a review", Data Mining and Knowledge Discovery, Vol. 33, No. 4, 917–963 (2019).

[19] Niu Z., Zhong G., Yu H., "A review on the attention mechanism of deep learning", *Neurocomputing*, Vol. 452, 48–62 (2021).

[20] Giantsidi S., Tarantola C., "Deep learning for financial forecasting: A review of recent trends", *International Review of Economics & Finance*, Vol. 104, 104719–104719 (2025).

[21] Kim J., Kim H., Kim H., Lee D., Yoon S., "A comprehensive survey of deep learning for time series forecasting: Architectural diversity and open challenges", *Artificial Intelligence Review*, Vol. 58, Art. 216 (2025).

[22] Cont R., "Empirical properties of asset returns: stylized facts and statistical issues", Quantitative Finance, Vol. 1, No. 2, 223–236 (2001).

[23] Heinlein R., Lepori G.M., "Do financial markets respond to macroeconomic surprises? Evidence from the UK", *Empirical Economics*, Vol. 62, 2329–2371 (2022).

[24] Singh V., Rocab E.D., "China's geopolitical risk and international financial markets: evidence from Canada", *Applied Economics*, Vol. 54, No. 34, 3953–3971 (2022).

Architektury głębokiego uczenia w prognozowaniu finansowych szeregów czasowych: porównawczy przegląd metod

M. CZYŻEWSKA

W artykule przedstawiono przegląd porównawczy wybranych architektur głębokiego uczenia stosowanych w prognozowaniu finansowych szeregów czasowych, ze szczególnym uwzględnieniem sieci Long Short-Term Memory (LSTM), konwolucyjnych sieci neuronowych (CNN) oraz hybrydowych modeli opartych na mechanizmie uwagi. Rozważania osadzono w kontekście trzech reprezentatywnych środowisk rynkowych: kryptowalut, indeksów giełdowych oraz rynku walutowego, aby wskazać, w jaki sposób różnice w zmienności, płynności i strukturze rynku mogą wpływać na dobór modelu. W pracy omówiono główne wyzwania prognozowania finansowego, w tym nieliniowość, heterogeniczność czasową oraz wrażliwość na zmieniające się warunki rynkowe. Szczególną uwagę poświęcono modelom hybrydowym i opartym na mechanizmie uwagi, zwłaszcza modelowi Regularized Self-Attention Regression (RSAR), jako przykładowi bardziej adaptacyjnego podejścia prognostycznego. Przegląd wskazuje, że żadna pojedyncza architektura nie może być uznana za uniwersalnie najlepszą dla wszystkich finansowych szeregów czasowych. Dobór modelu powinien zależeć od właściwości analizowanych danych, horyzontu prognozy oraz ograniczeń praktycznych, takich jak złożoność obliczeniowa i wrażliwość na dobór hiperparametrów.

Słowa kluczowe: prognozowanie finansowe, głębokie uczenie, modele hybrydowe.