SCIREA Journal of Computer
ISSN: 2995-6927
http://www.scirea.org/journal/Computer
May 29, 2026
Volume 11, Issue 1, February 2026
https://doi.org/10.54647/computer520478

Investigating the Effectiveness of Hybrid Deep Learning Networks Based on Self-Organization in Forecasting the Dynamics of Financial Markets

Yuriy Zaychenko[1], **Yevgeniy Bodyanskiy**[2], **Helen Zaichenko**[1], **and Oleksii Kuzmenko**[1,*]

[1]Educational and Research Institute for Applied System Analysis, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", 37-A Beresteiskyi Ave., Kyiv, 03056, Ukraine

[2]Artificial Intelligence Department, Kharkiv National University of Radio Electronics, 14 Nauky Ave., Kharkiv, 61166, Ukraine

***Corresponding author:** Oleksii Kuzmenko, oleksii.kuzmenko@ukr.net

Abstract

The article investigates the effectiveness of applying hybrid deep learning networks based on self-organization to the problem of forecasting financial market dynamics. The focus is placed on the GMDH-Neo-Fuzzy architecture, which combines the advantages of the Group Method of Data Handling (GMDH) and Neo-Fuzzy Neurons. The feasibility of using such models for the analysis of non-stationary and nonlinear time series, which are characteristic of modern financial markets, is substantiated.

The experimental study was conducted using the time series of daily closing prices of NVIDIA Corporation (NVDA) stock for the period 2024-2026. To evaluate the effectiveness of the proposed approach, a comparison of three forecasting models was performed: classical GMDH, GMDH-Neo-Fuzzy, and LSTM. The evaluation was carried out for short-term and middle-term forecasting horizons (1, 3, 5, 7, 14, 21, and 28 days) using the metrics MSE, MAPE, and training time.

The obtained results showed that the GMDH-Neo-Fuzzy model provides the best balance between forecasting accuracy and computational efficiency. In particular, it demonstrated the best performance for most forecasting horizons, especially for intervals from 3 to 28 days, significantly outperforming classical GMDH and, in most cases, LSTM. In addition, it was found that the training time of GMDH-Neo-Fuzzy is orders of magnitude lower compared to the recurrent LSTM network.

It is concluded that hybrid deep learning networks based on self-organization are a promising tool for building intelligent forecasting systems for financial time series.

Keywords: hybrid networks, deep learning, self-organization, forecasting, time series.

1. Introduction

The current state of global financial markets is characterized by an unprecedented level of dynamism, volatility, and unpredictability. In the context of globalization and digitalization of the economy, market indicators become extremely sensitive to a wide range of factors – from macroeconomic shifts to sentiment in social networks. This issue is particularly acute in the technology sector, where the market capitalization of leading companies such as NVIDIA Corporation (NVDA) is subject to significant fluctuations within just a few trading sessions [1]. Such conditions impose stringent requirements on the accuracy and performance of decision support systems.

Traditional approaches to the analysis and forecasting of time series, based on classical statistical methods such as autoregressive and moving average models (ARMA, ARIMA) or heteroskedastic models (ARCH, GARCH), often prove insufficient in modern realities. The main reason lies in their reliance on assumptions of linearity and stationarity of processes, which fundamentally contradicts the nonlinear and non-stationary nature of financial data.

A significant breakthrough in this field is associated with the development of deep learning methods [2]. In particular, Recurrent Neural Networks (RNNs) and their advanced modification – Long Short-Term Memory (LSTM) – have demonstrated the ability to capture complex long-term dependencies in data [3]. However, the practical implementation of LSTM in real-time trading systems is accompanied by a number of critical drawbacks: the complexity of selecting an optimal architecture, long training times, the risk of overfitting on limited datasets, and the "black box" problem, which complicates the interpretation of the obtained results.

In the context of the evolution of artificial neural networks across different generations, there arises a need to search for methods that combine the approximation capabilities of neural networks with the transparency and efficiency of self-organization algorithms [4]. One of the most promising directions is the use of hybrid computational intelligence systems.

This paper considers the GMDH-Neo-Fuzzy architecture, which integrates the Group Method of Data Handling (GMDH) and the properties of Neo-Fuzzy Neurons. The advantage of such networks lies in their ability for automatic structural and parametric adaptation (self-organization), which makes it possible to minimize the subjective influence of the researcher on model construction [5].

The main hypothesis is that the use of hybrid deep learning networks based on self-organization allows not only to achieve a significant advantage in computational speed but also providing higher forecasting accuracy compared to massive recurrent neural network architectures.

2. Architecture of Hybrid Deep Learning Network Based on Self-Organization

The hybrid deep learning network GMDH-Neo-Fuzzy, whose architecture is presented in Fig. 1, combines the advantages of inductive modeling based on the Group Method of Data Handling (GMDH) and the flexibility of fuzzy neural networks.

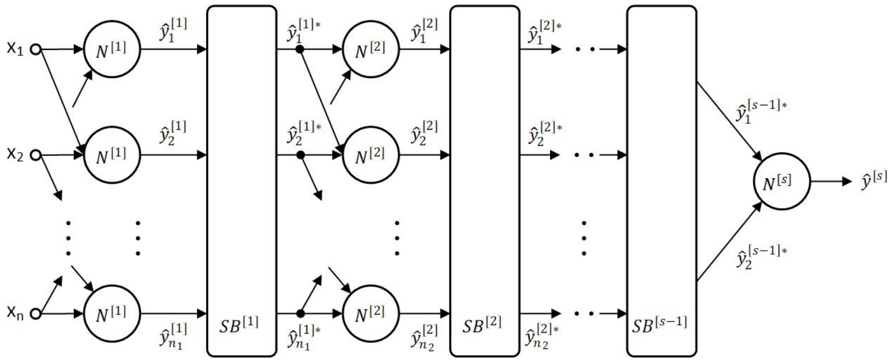

Figure 1. Architectural diagram of the GMDH-Neo-Fuzzy network

The first layer of the network plays the role of a receptor layer, i.e., it receives an n-

dimensional vector of input signals $x_1, x_2, ..., x_n$. As an example, Fig. 2 illustrates the process of network synthesis with the number of input signals $n = 4$.

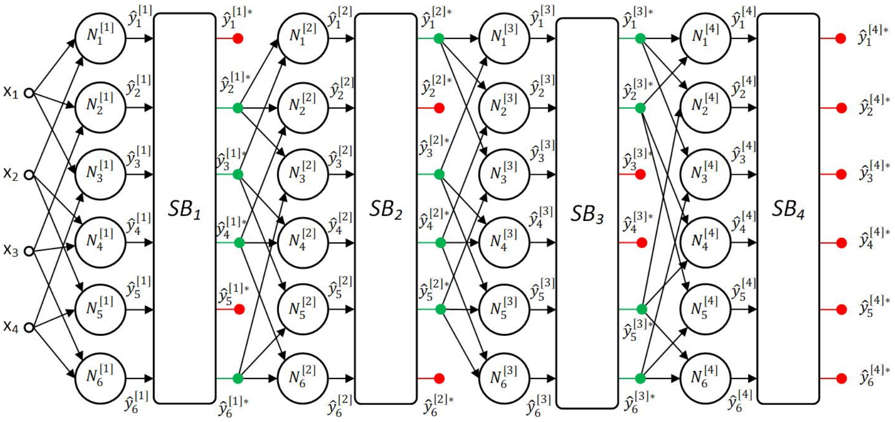

Figure 2. The process of synthesizing the structure of the GMDH- Neo- Fuzzy network

After passing through the receptor layer, the signal vector is fed to the first hidden layer, which contains nodes $N^{[1]}$ with two inputs. Their number is determined by formula $n_1 = C_n^2$ (for this case $n_1 = 6$). The first layer forms output signals $\hat{y}_l^{[1]}$, where $l = 1, 2, ..., 0.5n(n-1)$, which are passed to the first selection block $SB^{[1]}$. It is designed to select the best variants $n_1^* < n_1$ from the set $\hat{y}_l^{[1]}$ based on the variance value $\sigma^2_{\hat{y}_l^{[1]}}$. Next, combinations $\hat{y}_l^{[1]*}$, $\hat{y}_p^{[1]*}$ are formed in the amount $n_2 = C^2_{n_1^*}$, where $(n \leq n_2 \leq 2n)$. The formed signals are fed to the second hidden layer with nodes $N^{[2]}$. Then, the signals $\hat{y}_l^{[2]}$, obtained after training the nodes of this hidden layer, are passed to the selection block $SB^{[2]}$, where the best variants are selected based on accuracy, which must exceed the accuracy of the best signal from the previous layer $\hat{y}_l^{[1]*}$. The process of forming signals in subsequent hidden layers proceeds similarly and continues until the signal formed at some selection block $SB^{[s+1]}$ becomes worse than the best signal at the previous layer s according to the condition $\sigma^2_{\hat{y}_l^{[s+1]}} > \sigma^2_{\hat{y}_l^{[s]}}$. In this case, this occurred after passing the fourth selection block $SB^{[4]}$.

At the next stage, the synthesis process proceeds in the backward direction (Fig. 3).

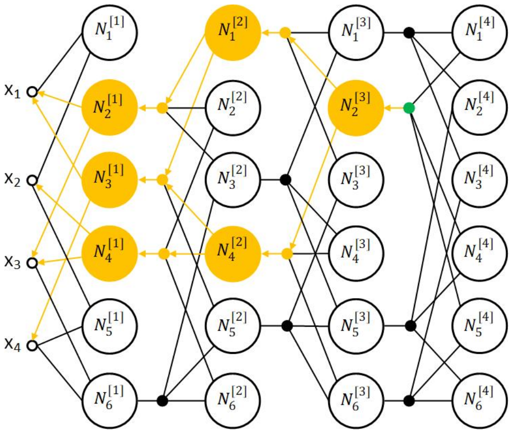

Figure 3. Backward direction of synthesizing the structure of the GMDH- Neo- Fuzzy network

First, the best node $N^{[s]}$ is determined, which is in the previous layer and has the output signal $\hat{y}^s$. Then, due to the known connections with the neurons of the previous hidden layer, a sequential traversal of all layers up to the initial one is performed [6]. As a result of this backward movement, the final structure of the GMDH- Neo- Fuzzy deep learning network is obtained (Fig. 4).

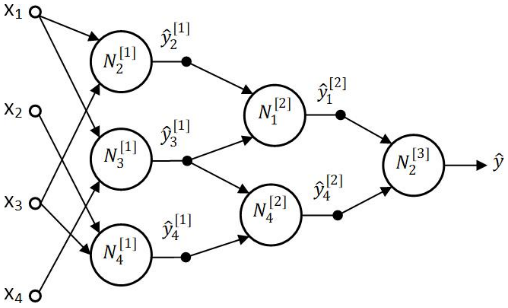

Figure 4. Result of synthesizing the structure of the GMDH- Neo- Fuzzy network

The approach considered makes it possible to obtain not only an optimal structure but also a well- trained network that can be immediately used for forecasting. Another advantage is the elimination of the vanishing gradient problem due to sequential layer- by- layer training. Since only a small number of neurons are trained, this allows the use of small training datasets and significantly reduces training time. The efficiency of training this hybrid network is explained by the fact that Neo- Fuzzy Neurons are used as nodes, which makes it necessary to adapt only the weight coefficients without modifying the parameters of membership functions.

Let us consider the architecture of a Neo-Fuzzy Neuron, presented in Fig. 5.

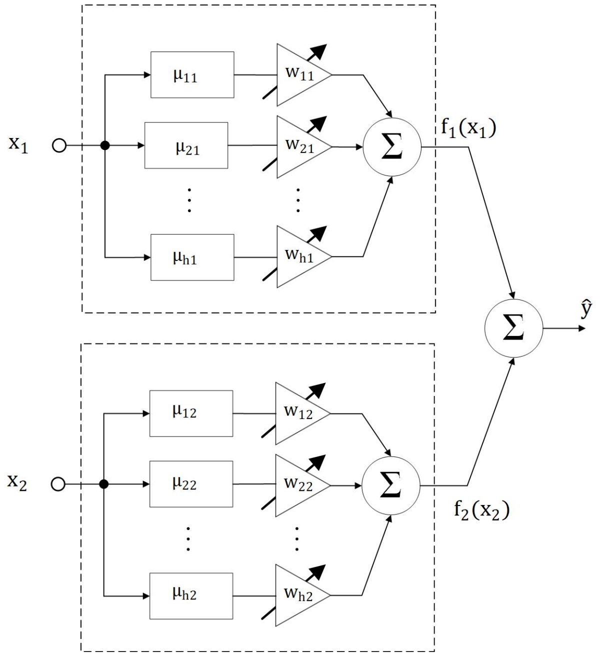

Figure 5. Architecture of a Neo- Fuzzy Neuron

The main difference of this node from the general structure of a Neo-Fuzzy Neuron [7] is that it has only two inputs ($n = 2$) and one output. It implements the following mapping:

$$\hat{y} = \sum_{i=1}^{n} f_i(x_i) \tag{1}$$

where x_i is the i-th input ($i = 1, 2, ..., n$), and $\hat{y}$ is the system output.

The structural elements of the Neo- Fuzzy Neuron are nonlinear synapses NS_i, which are indicated by dashed rectangles in the diagram. Their role is to transform the input signal:

$$f_i(x_i) = \sum_{j=1}^{h} w_{ji}\mu_{ji}(x_i) \tag{2}$$

and to implement fuzzy inference: if x_i is x_{ji}, then the output equals w_{ji}, where x_{ji} is a fuzzy set with membership function μ_{ji}, and w_{ji} is the synaptic weight in the consequent.

The learning criterion (objective function) is the standard local quadratic error function:

$$E(k) = \frac{1}{2}\big(y(k) - \hat{y}(k)\big)^2 = \frac{1}{2}e(k)^2 = \frac{1}{2}(y(k) - \sum_{i=1}^{2}\sum_{j=1}^{h} w_{ji}\mu_{ji}(x_i(k)))^2 \tag{3}$$

It is minimized using the traditional stochastic gradient descent algorithm. In the case where a priori defined dataset is available, the learning process can be performed in batch mode in a single epoch using the least squares method:

$$w^{[1]}(N) = \left(\sum_{k=1}^{N} \mu^{[1]}(k)\mu^{[1]T}(k)\right)^{+} \sum_{k=1}^{N} \mu^{[1]}(k)y(k) = P^{[1]}(N)\sum_{k=1}^{N} \mu^{[1]}(k)y(k) \tag{4}$$

where $(\bullet)^{+}$ denotes the Moore–Penrose pseudoinverse (here $y(k)$ denotes the external reference signal). If training observations are provided sequentially in real time (online), a recursive least squares (RLS) form can be used:

$$\begin{cases} w_l^{ij}(k) = w_l^{ij}(k-1) + \dfrac{P^{ij}(k-1)(y(k) - \left(w_l^{ij}(k-1)\right)^T \varphi^{ij}(x(k)))\varphi^{ij}(x(k))}{1 + \left(\varphi^{ij}(x(k))\right)^T P^{ij}(k-1)\varphi^{ij}(x(k))} \\ P^{ij}(k) = P^{ij}(k-1) - \dfrac{P^{ij}(k-1)\varphi^{ij}(x(k))\left(\varphi^{ij}(x(k))\right)^T P^{ij}(k-1)}{1 + \left(\varphi^{ij}(x(k))\right)^T P^{ij}(k-1)\varphi^{ij}(x(k))} \end{cases} \tag{5}$$

3. Description and Analysis of the Dataset

The process of studying financial markets and developing forecasting systems for highly volatile asset prices requires thorough preparation and analysis of the source data. In this study, the object of analysis was the time series of daily closing prices of NVIDIA Corporation (NVDA) stock, which, as of 2024-2026, holds a leading position in the segment of high-performance computing and artificial intelligence infrastructure. The choice of this particular asset is determined by its strategic importance for the global economy and the growth of its market capitalization, which is accompanied by complex dynamic processes that are difficult to forecast using classical statistical methods [8, 9, 10].

The source data for the experiments were obtained from the Yahoo Finance platform [1]. The dataset covers the period from February 12, 2024, to February 9, 2026, comprising 502 trading sessions. Each record in the sample contains the date and the corresponding closing price in US dollars. The use of closing prices instead of intraday fluctuations makes it possible to focus on the resulting market valuation of the assets at the end of the trading day, smoothing out random noise.

The dynamics of NVDA prices over the specified period reflect the rapid development of the artificial intelligence ecosystem. The initial price in February 2024 was approximately 72.25 USD, but after mid-2025 it consistently exceeded the 150 USD mark, and by the beginning of

2026 it approached the 190 USD level. Such growth was driven by record data center revenues and demand for the Blackwell and Rubin architecture.

Figure 6 clearly shows a long-term upward trend; however, the analysis also reveals periods of significant correction, particularly in April 2025, when the price temporarily dropped below 100 USD due to a general market correction and a revision of investor expectations. Such fluctuations make this dataset suitable for evaluating adaptive forecasting models such as GMDH- Neo- Fuzzy.

Figure 6. Dynamics of NVIDIA Corporation (NVDA) stock closing prices during the period 2024- 2026

One of the most important stages in time series analysis is checking for stationarity, since non-stationarity is a typical characteristic of financial time series and significantly affects the choice of forecasting model [11, 8]. A stationary series has a constant mean, variance, and autocorrelation structure over time. To test for the presence of a unit root in the NVDA series, the Augmented Dickey- Fuller (ADF) test was applied, which is one of the basic tools for analyzing time series stationarity in financial research [11].

The results of the ADF test for the NVDA price level are presented in Table 1.

Table 1. Results of the Dickey- Fuller (ADF) test for the NVDA price series

Indicator	**Value**
ADF Statistic	-1.45511
p-value	0.55559
Critical Value (1%)	-3.44363

Critical Value (5%)	-2.86740
Critical Value (10%)	-2.56989

Since the obtained ADF statistic value (-1.45511) is greater (less negative) than the critical values at all standard significance levels, and the p-value is 0.55559, we have no grounds to reject the null hypothesis of the presence of a unit root. This mathematically confirms that the NVDA price time series is non- stationary. Non- stationarity indicates that the series contains a trend component or stochastic drift, which is characteristic of financial assets whose prices are shaped by new information flows, changes in market regimes, and dynamic volatility [11, 10]. For practical modeling, this implies the need to use methods capable of working with non-stationary data or to apply differencing to the series.

To identify the internal memory of the time series, an analysis of the autocorrelation function (ACF) and the partial autocorrelation function (PACF) was carried out. These functions make it possible to determine the extent to which the current price value depends on its previous states. Figure 7 shows the ACF plot, which demonstrates a very slow decay in lag significance. The autocorrelation coefficient for the first lag is 0.99, and even at lag 50 it remains at 0.74. Such a structure is a classical indicator of non-stationarity and indicates the presence of a strong long- term dependence in the data. The NVIDIA stock market demonstrates high inertia: yesterday's price is a very accurate predictor of today's price, which explains the success of short- term forecasts.

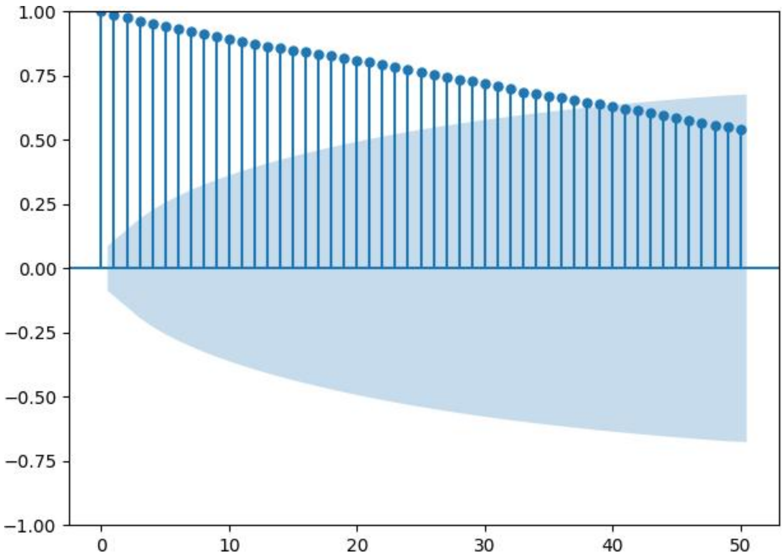

Figure 7. Correlogram (ACF) of NVDA Close Price

The partial autocorrelation function (PACF) in Fig. 8, by contrast, typically shows a significant peak only at the first lag, indicating that after accounting for the influence of the first lag, the correlation at subsequent steps becomes negligible. This makes the use of autoregressive (AR) models or LSTM networks, which focus on the dynamics of short-term changes, appropriate.

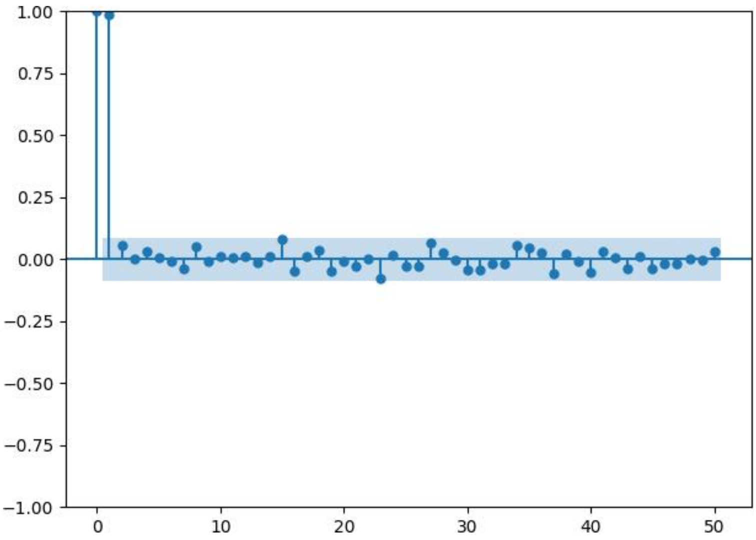

Figure 8. Partial autocorrelation function (PACF) of NVDA Close Price

To verify the randomness of the data and the presence of hidden patterns, a lag plot was used, presented in Fig. 9. It is characterized by a step of $k = 1$, and the values Y_t are plotted relative to Y_{t-1}. The points in the plot are densely grouped along the diagonal line, which confirms the absence of white noise and the presence of a clear deterministic structure in the dynamics of NVDA stock. Such a linear structure indicates the appropriateness of using autoregressive models and machine learning methods for identifying patterns.

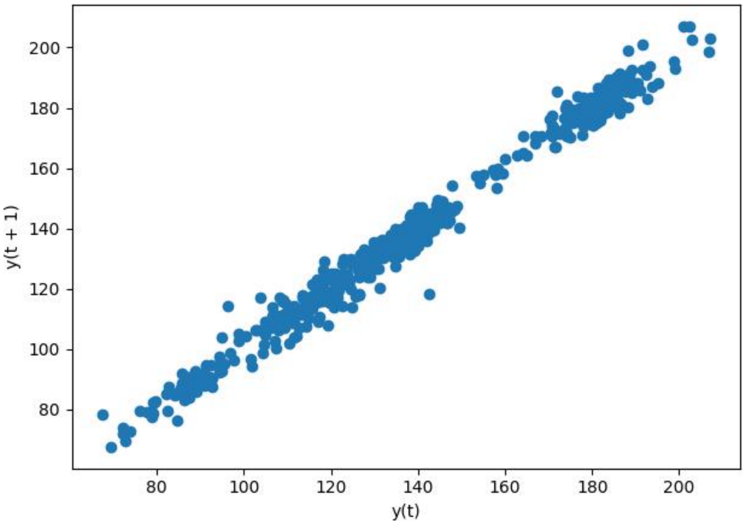

Figure 9. Lag plot for NVDA Close Price

The analysis of the distribution density (Fig. 10) shows that closing prices do not follow normal distribution. Several local maxima are observed, corresponding to periods of price stabilization at certain levels during trend formation. The wide distribution range from 70 to 200 USD reflects the large-scale market revaluation of the company over the studied period.

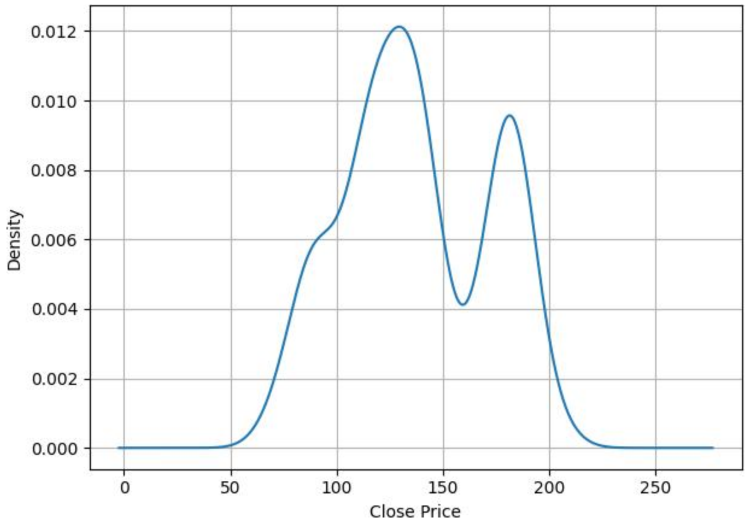

Figure 10. Kernel density estimation (KDE) of the NVDA price distribution

Thus, the analysis of the NVDA dataset confirms its complex and dynamic nature, where a strong trend is combined with periods of volatility, making it a suitable testbed for evaluating modern hybrid deep learning network models [8, 12].

4. Experimental Approach

The process of organizing experiments within this study is based on a comprehensive approach to the analysis and forecasting of volatile financial time series, in particular the stock price of NVIDIA Corporation (NVDA). Given the complex nature of the stock market, where price fluctuations are determined both by fundamental economic factors and stochastic noise, the selection of intelligent forecasting methods focused primarily on adaptive approaches and deep learning networks [13]. The aim of the experimental part is to test the hypothesis that hybrid deep learning networks based on self-organization are no less effective in terms of accuracy and training speed compared to other deep learning models.

The experiments were conducted across three classes of models, each characterized by specific theoretical foundations, which makes it possible to cover a wide range of dynamic characteristics of the time series.

First, the Group Method of Data Handling (GMDH) was selected [14, 15, 16]. This is a representative of the inductive approach, where the model structure is not specified a priori but is synthesized automatically based on the training data. GMDH uses the principle of external supplementation, which makes it possible to avoid overfitting by testing partial models on an independent validation set. The choice of GMDH is justified by its ability to work with relatively short datasets and to automatically select the most informative lag variables.

Second, a hybrid deep learning network based on self-organization (GMDH-Neo-Fuzzy) was applied. This model combines the structural flexibility of GMDH with the interpretability and ability to approximate uncertain relationships inherent in fuzzy systems. Neo-Fuzzy Neurons make it possible to effectively process noise in financial data and take into account the uncertainty of market fluctuations, which is consistent with modern studies on hybrid and fuzzy-neural forecasting architectures [17]. This is critically important for forecasting NVIDIA stock, since the technology sector is often subject to the influence of irrational market expectations.

Third, Long Short-Term Memory (LSTM) was used. As a specialized form of recurrent neural networks, LSTM possesses a gating mechanism (input, forget, and output gates), which allows it to retain information about long-term dependencies and significantly mitigate the vanishing gradient problem [18, 19, 20]. In financial forecasting, LSTM is effective for identifying complex temporal patterns in large datasets [20, 19].

The series of experiments were organized according to a multiskilling principle, where forecasting was performed for seven different time intervals (forecast horizons): 1, 3, 5, 7, 14, 21, and 28 days. This approach makes it possible to evaluate the degradation of model accuracy as the forecast depth increases, which is a typical feature of multi-step time series forecasting [21, 22, 23].

For each interval, the optimal window size (the number of input lag values) was determined based on the autocorrelation function of the time series.

The data were split into training and validation sets according to percentage ratios. The following variations were adopted: 80/20, 70/30, and 60/40. In addition, a separate test dataset containing 30 days of observations was allocated for model evaluation. It was not used for training or validation but was used exclusively for assessing and comparing all models.

To quantitatively assess forecast quality and compare model performance, a system of metrics was selected (Table 2), making it possible to evaluate both forecasting accuracy and computational cost.

Table 2. Metrics for forecasting quality evaluation

Metric	**Description**	**Features**
MSE (Mean Squared Error)	Reflects the variance of the forecast relative to the actual values.	Sensitive to large deviations, which is important when assessing risks in the stock market.
MAPE (Mean Absolute Percentage Error)	Provides an intuitively understandable estimate of forecast accuracy in percentage terms.	Sensitive to small actual values.
Training Time	Reflects the duration of model training.	A critical parameter for real-time systems; allows evaluation of the feasibility of using complex models.

The organization of the experiments assumes that each model undergoes a cycle of training and testing independently for each of the seven intervals. This makes it possible to form a complete picture of performance: from the ability to capture immediate volatility to forecast stable middle-term trends, which corresponds to modern approaches to multi-step forecasting, where model quality should be evaluated separately for each forecast horizon [24, 25].

5. Experimental Results and Analysis

During the study, experimental results were obtained for forecasting the stock price of NVIDIA Corporation (NVDA) using three architectures: GMDH, GMDH-Neo-Fuzzy, and LSTM. The experiments were conducted in two stages: short-term forecasting (with intervals of 1, 3, 5, and 7 days) and middle-term forecasting (with intervals of 14, 21, and 28 days).

5.1 Short-Term Forecasting

Short-term forecasting is the most critical scenario for applied financial systems, since it is precisely on short horizons (1-7 trading days) that the basis for prompt decision-making is formed in tasks such as algorithmic trading, portfolio rebalancing, and short-term risk management.

For the 1-day horizon, all three models demonstrated a high level of accuracy, which is expected for such a short forecasting horizon, where the inertia of the series is still preserved. Classical GMDH provided the best values of mean squared error and relative error: $MSE = 15.73$ and $MAPE = 1.56\%$. At the same time, LSTM showed an almost comparable result ($MSE = 16.46$, $MAPE = 1.59\%$), while GMDH-Neo-Fuzzy was only slightly inferior in terms of accuracy ($MSE = 16.57$, $MAPE = 1.66\%$) but significantly outperformed the others in training speed.

To illustrate the behavior of the best-performing GMDH model for the 1-day horizon, the forecast plot is presented in Fig. 11. Analysis of the plot in Fig. 11 shows that the GMDH model reproduces the local price trajectory quite accurately; however, deviations occur at points of sharp fluctuations. In particular, for certain dates within the test interval, noticeably larger local errors are observed, which is typical of financial time series with elevated volatility. This means that although the model reproduces the overall dynamics well, it is less sensitive to short impulsive changes.

It is also important to note that the training time of the GMDH-Neo-Fuzzy model in this case was only 15 ms, whereas for LSTM it exceeded 27 s. This indicates that even with a slight loss in accuracy, the hybrid network demonstrates substantially better computational characteristics, which is critically important for real-time systems.

Figure 11. Results of short-term 1-day forecasting for the GMDH model

At the 3-day horizon, the situation changed fundamentally. While classical GMDH and LSTM showed a deterioration in forecast quality, GMDH-Neo-Fuzzy, on the contrary, sharply improved its results and became the undisputed leader among all models.

According to the experiments, for the 3-day horizon the GMDH-Neo-Fuzzy model achieved $MSE = 12.55$ and $MAPE = 1.41\%$, which is the best result among all the studied architectures. For comparison, classical GMDH obtained $MSE = 37.64$ and $MAPE = 2.46\%$, while LSTM yielded $MSE = 31.77$ and $MAPE = 2.22\%$.

To illustrate this result, the forecast plot of the GMDH-Neo-Fuzzy network with a 3-day interval is shown in Fig. 12.

Figure 12. Results of short-term forecasting with a 3-day interval for the GMDH-Neo-Fuzzy model

This result can be explained by the fact that at the 3-day horizon the financial series already requires not only local approximation of neighboring values, but also the detection of short nonlinear relationships between lags. It is precisely here that the advantage of fuzzy neurons becomes most evident: they make it possible to combine the interpretability of the self-organized GMDH structure with the flexibility of fuzzy nonlinear mapping.

To illustrate not only the forecasting results but also the model structure, Fig. 13 shows the synthesized architecture of the GMDH-Neo-Fuzzy network for a 3-day horizon.

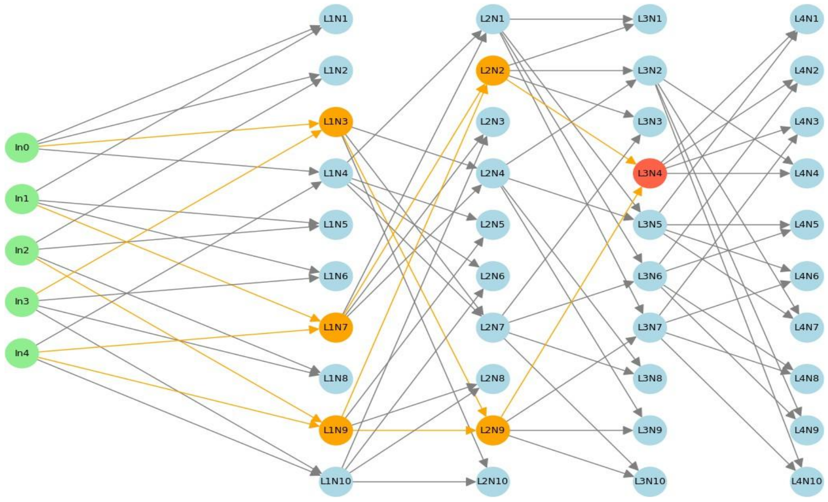

Figure 13. Synthesized architecture of the GMDH-Neo-Fuzzy model for short-term forecasting with a 3-day horizon

As can be seen from Fig. 13, the model starts with five inputs $In0, \ldots, In4$, which corresponds to lagged values of the financial time series. At the first layer, a set of candidate nodes $L1N1, \ldots, L1N10$ is formed; however, only a subset of them passes the first selection block. Subsequently, at the second and third layers, the self-organization mechanism again selects only the most relevant nodes. These elements capture the most significant nonlinear dependencies between the lagged values of the series and provide the formation of informative intermediate representations. At the fourth layer, the synthesis process stops, and at the previous layer the output node $L3N4$ is determined, which serves as the final aggregator of all essential features identified during the multi-level self-organization process.

For the 5-day horizon, the superiority of GMDH-Neo-Fuzzy was preserved. According to the experimental results, this model achieved $MSE = 16.54$ and $MAPE = 1.71\%$, which is significantly better than classical GMDH ($MSE = 46.87$, $MAPE = 2.69\%$) and LSTM ($MSE = 39.00$, $MAPE = 2.38\%$). This indicates that at the 5-day horizon the market still retains a relatively compact short-term pattern but already requires a more flexible nonlinear model than classical GMDH. At the same time, although LSTM is capable of reproducing the overall shape of the series, it proves less effective due to its higher inertia and training complexity.

At the 7-day horizon, the task became even more difficult, since the forecast error accumulates and the influence of local noise and nonlinear market factors increases. Nevertheless, even in this case GMDH-Neo-Fuzzy retains the best results: $MSE = 15.70$ and $MAPE = 1.61\%$. For comparison, GMDH has $MSE = 51.57$ and $MAPE = 2.82\%$, while LSTM has $MSE = 32.19$ and $MAPE = 2.52\%$.

The summarized results of short-term forecasting are presented in Table 3.

Table 3. Short-term forecasting metrics

Interval	Model	MSE	MAPE (%)	Training Time (ms)
1 day	GMDH	**15.73**	**1.56**	489.0
	GMDH-Neo-Fuzzy	16.57	1.66	**15.0**
	LSTM	16.46	1.59	27386.0
3 days	GMDH	37.64	2.46	205.0
	GMDH-Neo-Fuzzy	**12.55**	**1.41**	**20.0**

	LSTM	31.77	2.22	7617.0
5 days	GMDH	46.87	2.69	366.0
	GMDH-Neo-Fuzzy	**16.54**	**1.71**	**10.0**
	LSTM	39.0	2.38	8058.0
7 days	GMDH	51.57	2.82	298.0
	GMDH-Neo-Fuzzy	**15.7**	**1.61**	**18.0**
	LSTM	32.19	2.52	5762.0

Analysis of Table 3 confirms a clear pattern: as the forecast horizon increases, the accuracy of classical GMDH systematically deteriorates, whereas GMDH-Neo-Fuzzy maintains consistently low MSE and MAPE values. On average, for intervals of 3-7 days, the hybrid model provides the best forecast quality, significantly outperforming both classical GMDH and LSTM. It should also be emphasized that the training time of GMDH-Neo-Fuzzy across all short-term horizons is only 10-20 ms, which is several orders of magnitude lower than that of LSTM.

5.2 Middle-Term Forecasting

Unlike short-term forecasting, in this case the models must not only capture local dependencies between neighboring observations, but also correctly reproduce longer trend and quasi-cyclic components of the time series.

As the forecast horizon increases, the approximation task naturally becomes more difficult, while the influence of random market fluctuations, short-term impulses, and nonlinear structures becomes more significant. For this reason, the analysis of middle-term intervals is particularly informative from the perspective of the generalization ability of the investigated models.

For the 14-day forecasting horizon, classical GMDH already demonstrates a noticeable deterioration in quality: $MSE = 58.22$ and $MAPE = 3.46\%$. In contrast, the two neural network architectures provided significantly better results. The best overall model in terms of the selected metrics was GMDH-Neo-Fuzzy, which achieved $MSE = 23.57$ and $MAPE = 1.83\%$, while LSTM showed a close but slightly weaker result: $MSE = 25.80$ and $MAPE = 1.84\%$.

This indicates that the hybrid model reproduces the overall direction of movement of the series and the key segments of the middle-term trend quite well. It is particularly important that even for middle-term intervals, GMDH-Neo-Fuzzy maintains a very low training time – only 18 ms, whereas for LSTM this indicator is nearly 6 s. This means that with practically the same level of accuracy, the hybrid self-organized architecture is significantly more attractive from the perspective of computational efficiency.

For the 21-day horizon, an even more pronounced advantage of the GMDH-Neo-Fuzzy model is observed. In this case, classical GMDH showed substantially worse results ($MSE = 101.33$, $MAPE = 5.05\%$), and although LSTM remained competitive ($MSE = 24.61$, $MAPE = 2.05\%$), it still underperformed relative to the hybrid model. Once again, the best result was demonstrated by GMDH-Neo-Fuzzy: $MSE = 18.50$ and $MAPE = 1.66\%$.

From a practical point of view, the 21-day horizon is one of the most interesting, since it approximately corresponds to one trading month.

For the 28-day horizon, a somewhat different picture is observed. As before, classical GMDH demonstrates the worst result among all models ($MSE = 227.42$, $MAPE = 7.46\%$), which indicates a substantial degradation in its accuracy as the forecasting horizon increases. However, at this interval the competition between GMDH-Neo-Fuzzy and LSTM becomes extremely close. According to the mean squared error criterion, GMDH-Neo-Fuzzy remains the best model ($MSE = 20.56$), whereas in terms of relative error, LSTM obtains a slight advantage ($MAPE = 1.72\%$ versus 1.74% for GMDH-Neo-Fuzzy).

The summarized results of middle-term forecasting are presented in Table 4.

Table 4. Middle-term forecasting metrics

Interval	**Model**	**MSE**	**MAPE (%)**	**Training Time (ms)**
14 days	GMDH	58.22	3.46	347.0
	GMDH-Neo-Fuzzy	**23.57**	**1.83**	**18.0**
	LSTM	25.8	1.84	5993.0
21 days	GMDH	101.33	5.05	405.0
	GMDH-Neo-Fuzzy	**18.5**	**1.66**	**30.0**
	LSTM	24.61	2.05	5693.0

28 days	GMDH	227.42	7.46	122.0
	GMDH-Neo-Fuzzy	**20.56**	1.74	**31.0**
	LSTM	22.54	**1.72**	5903.0

Analysis of Table 4 demonstrates a clear trend: as the forecasting horizon increases, the accuracy of classical GMDH deteriorates sharply, whereas GMDH-Neo-Fuzzy maintains a consistently high level of quality. For the 14- and 21-day intervals, this model is the undisputed leader both in terms of MSE and $MAPE$. At the 28-day horizon, LSTM only slightly outperforms GMDH-Neo-Fuzzy in terms of $MAPE$, but still lags behind it in terms of MSE and, most importantly, is substantially inferior in training time.

5.3 Comparison of Models by Forecasting Accuracy

To summarize the results of all conducted experiments, it is appropriate to perform a comparative analysis of the accuracy of the investigated models across all forecasting horizons. Figures 14 and 15 present a comparison of the GMDH, GMDH-Neo-Fuzzy, and LSTM models using two key metrics – mean squared error (MSE) and mean absolute percentage error (MAPE).

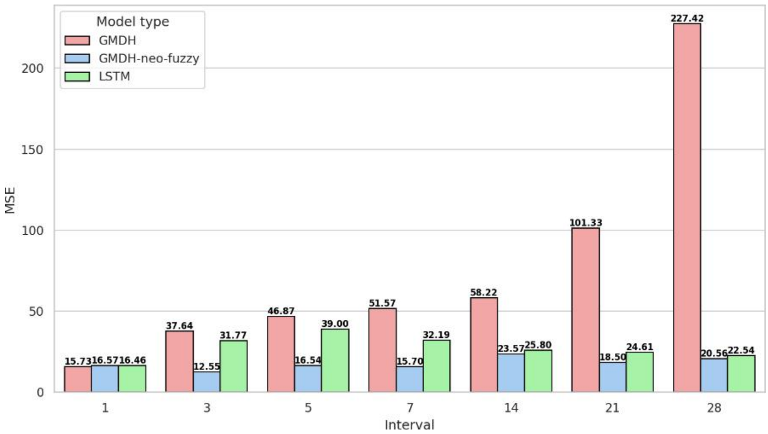

Figure 14. Comparison of models by the MSE metric at different forecasting intervals

Analysis of the graph in Fig. 14 demonstrates the most general and illustrative pattern in model behavior: as the forecasting horizon increases, the accuracy of classical GMDH deteriorates rapidly. While at the 1-day horizon the MSE value for GMDH is only 15.73 and is the best among all models, by the 7-day interval it increases to 51.57, and by the 28-day horizon it reaches 227.42. Such dynamics indicate that classical GMDH is effective only for short forecasts, when the local inertia of the series dominates, but loses adequacy when more complex middle-term dependencies must be reproduced.

In contrast to classical GMDH, the GMDH-Neo-Fuzzy model demonstrates the most stable behavior across the entire studied range of forecasting horizons. Its MSE values remain within the range from 12.55 to 23.57 for most intervals. It is particularly noteworthy that GMDH-Neo-Fuzzy provides the best MSE result at the 3-, 5-, 7-, 14-, 21-, and 28-day horizons, outperforming all other models except for the 1-day forecast. This means that the hybrid self-organized architecture is the most robust to the accumulation of forecasting errors.

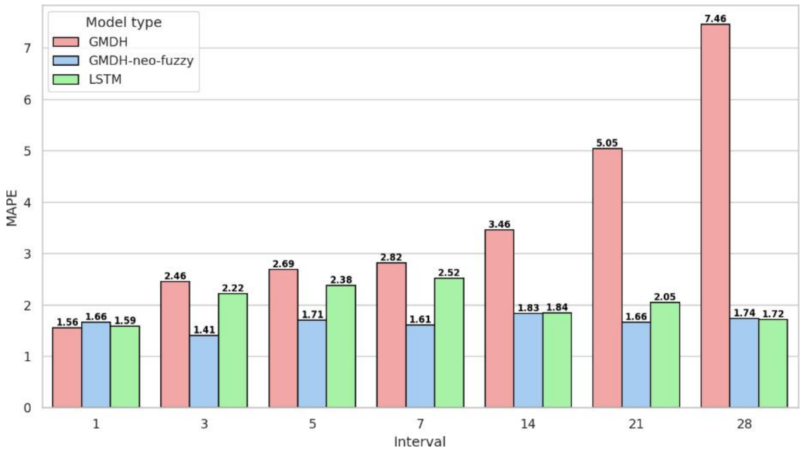

Figure 15. Comparison of models by the MAPE metric at different forecasting intervals

Thus, the comparative analysis of the MSE and $MAPE$ metrics confirms that the hybrid GMDH-Neo-Fuzzy model has the highest generalization ability among all the studied architectures. It combines high accuracy at both short-term and middle-term horizons while maintaining robustness to the increasing complexity of the forecasting task. This is precisely what allows it to be considered the most suitable model for building practical forecasting systems for financial market dynamics.

5.4 Comparison of Models by Training Time

In addition to forecasting accuracy, an important criterion of a model's practical applicability is its computational efficiency, particularly the training time. For financial analytics and decision support systems operating in near real time, this indicator is of fundamental importance, since it determines the possibility of promptly retraining the model on new data.

The average training times for the studied models are presented in Table 5.

Table 5. Average model training time

Model	**Average Training Time (ms)**	**Relative Speed**
GMDH-Neo-Fuzzy	**20.29**	1x

GMDH	318.86	15.7x slower
LSTM	9487.43	467.6x slower

As can be seen from Table 5, the shortest average training time is achieved by the GMDH-Neo-Fuzzy model, which on average requires only 20.29 ms to build the model. Classical GMDH is approximately 15.7 times slower, while LSTM performs substantially worse in this regard – its training is on average almost 468 times slower.

The obtained results indicate that GMDH-Neo-Fuzzy has an indisputable advantage from the point of view of computational complexity. This makes it the most suitable for practical use in financial market forecasting tasks, where it is important to combine high accuracy with minimal time expenditure on training.

5.5 Advantages and Disadvantages of Models

Each of the studied models has its own strengths and weaknesses, which determine the appropriateness of its use depending on the nature of the forecasting task.

GMDH is distinguished by its relative simplicity of implementation, interpretability, and high efficiency on short forecasting horizons. At the same time, its main drawback is the rapid decline in accuracy as the forecasting horizon increases, which limits its suitability for complex nonlinear time series.

LSTM has the advantage of being able to identify long-term temporal dependencies and work with complex sequential data structures. For this reason, LSTM model demonstrates stable results in both short-term and middle-term forecasting. However, its use is accompanied by significant computational costs, long training times, and lower interpretability compared to self-organizing models.

The most balanced among the studied models proved to be GMDH-Neo-Fuzzy, which combines high accuracy, structural adaptability, fast training, and relative transparency of model construction. Its main advantage lies in its ability to automatically adapt its structure to the complexity of the data, which makes it possible to achieve consistently high results across different forecasting horizons. Possible disadvantages of this architecture include a more complex theoretical formulation compared to classical GMDH.

Thus, from a practical point of view, GMDH-Neo-Fuzzy is the most promising model for forecasting financial market dynamics, since it provides the best compromise between accuracy, speed, and adaptability.

6. Conclusions

As a result of the conducted study, the following main scientific and practical results were obtained:

1. An analysis of modern approaches to financial time series forecasting was carried out, and the feasibility of using hybrid deep learning networks based on self-organization for solving forecasting problems of non-stationary and nonlinear financial processes was substantiated.
2. A series of experiments on short-term and middle-term forecasting was conducted for horizons of 1, 3, 5, 7, 14, 21, and 28 days using three models: GMDH, GMDH-Neo-Fuzzy, and LSTM. The evaluation of forecasting quality was performed using the MSE, MAPE, and training time metrics.
3. It was proven that the hybrid GMDH-Neo-Fuzzy model is the most effective among the studied architectures in terms of the overall set of performance indicators. It demonstrated the best results in terms of MSE and MAPE across most forecasting horizons (3-28 days), and also showed the shortest training time among all the considered models.
4. It was determined that the main advantage of GMDH-Neo-Fuzzy lies in the successful combination of high forecasting accuracy, structural adaptability, robustness to increasing forecast horizons, and extremely low computational complexity. This makes it particularly promising for use in practical financial analytics and decision support systems operating in near real time.
5. The obtained results confirmed the main hypothesis of the study: the use of hybrid deep learning networks based on self-organization makes it possible to achieve not only high forecasting accuracy but also a significant advantage in computational efficiency compared to traditional and recurrent neural network architectures.

Thus, the use of the GMDH-Neo-Fuzzy hybrid network for forecasting financial market dynamics is both scientifically justified and practically feasible, and the model itself can be considered an effective foundation for building intelligent forecasting systems for stock market assets.

7. Further Research

Further research may be aimed at expanding the capabilities of the GMDH-Neo-Fuzzy model and increasing its practical value in financial market forecasting tasks. The main promising directions include:

- investigating the possibilities of automatic optimization of network parameters in order to improve its adaptability;
- testing the model on other financial instruments, in particular stock indices, currency pairs, and cryptocurrencies;
- developing an online version of the model for use in real-time forecasting systems.

Thus, the further development of the GMDH-Neo-Fuzzy model has significant potential for building effective intelligent systems for the analysis and forecasting of financial time series.

References

[1] Yahoo Finance. Nvidia corporation (NVDA) stock price, news, quote & history. https://finance.yahoo.com/quote/NVDA, 2026.

[2] T. Hall et al. A survey of machine learning methods for time series prediction. Applied Sciences, 15(11):5957, 2025.

[3] Y. Zaychenko and O. Kuzmenko. Investigation of hybrid deep learning networks and LSTM in the short term forecasting in financial sphere. In *2022 IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT)*, pages 65-68, Lviv, Ukraine, 2022.

[4] Y. Bodyanskiy, Y. Zaychenko, H. Zaichenko, and O. Kuzmenko. Investigation of the effectiveness of artificial neural networks of different generations in the task of forecasting in the financial sphere. *System Research and Information Technologies*, 1:124-137, 2025.

[5] Ye. Bodyanskiy, Yu. Zaychenko, O. Kuzmenko, and He. Zaichenko. Investigation of the effectiveness of the hybrid system of computational intelligence based on bagging and Group Method of Data Handling in the task of stock index forecasting. In *Proceedings of the IV International Scientific Symposium "Intelligent Solutions" (IntSol-2025)*, Uzhhorod, Ukraine, 2025. 1-5 May.

[6] Yuriy Zaychenko, Helen Zaichenko, and Galib Hamidov. Hybrid GMDH deep learning networks – analysis, optimization and applications in forecasting at financial sphere.

System Research and Information Technologies, 1:73–86, 2022.

[7] Takeshi Yamakawa, E. Uchino, T. Miki, and H. Kusabagi. A neo fuzzy neuron and its applications to system identification and predictions to system behavior. In *Proceedings of the International Conference*, pages 477–484, 1992.

[8] Zineb Bousbaa, Javier Sanchez-Medina, and Omar Bencharef. Financial time series forecasting: A data stream mining- based system. *Electronics*, 12(9):2039, 2023.

[9] Pavlos I. Zitis, Stelios M. Potirakis, and Alex Alexandridis. Forecasting forex market volatility using deep learning models and complexity measures. *Journal of Risk and Financial Management*, 17(12):557, 2024.

[10] Zhiang Qiu, Clemens Kownatzki, Fabien Scalzo, and Eun Sang Cha. Historical perspectives in volatility forecasting methods with machine learning. *Risks*, 13(5):98, 2025.

[11] Tolga Omay, Aysegul Corakci, and Esra Hasdemir. High persistence and nonlinear behavior in financial variables: A more powerful unit root testing in the ESTAR framework. *Mathematics*, 9(20):2534, 2021.

[12] Yung-Chen Li, Hsiao-Yun Huang, Nan-Ping Yang, and Yi-Hung Kung. Stock market forecasting based on spatiotemporal deep learning. *Entropy*, 25(9):1326, 2023.

[13] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.

[14] A. G. Ivakhnenko. Heuristic self- organization in problems of engineering cybernetics. *Automatica*, 6(2):207- 219, 1970.

[15] A. G. Ivakhnenko and I. A. Mueller. *Self- Organization of Forecasting Models*. Technika, Kiev, 1985.

[16] A. G. Ivakhnenko, G. A. Ivakhnenko, and J. A. Mueller. Self- organization of the neural networks with active neurons. *Pattern Recognition and Image Analysis*, 4(2):177- 188, 1994.

[17] Weina Wang, Jiapeng Shao, and Huxidan Jumahong. Fuzzy inference- based LSTM for long- term time series prediction. *Scientific Reports*, 13:20359, 2023.

[18] S. Hochreiter and J. Schmidhuber. Long short- term memory. *Neural Computation*, 9(8):1735- 1780, 1997.

[19] Song, Xiaobao & Deng, Liwei & Hao, Wang & Zhang, Yaoan & Yuxin, He & Cao, Wenming. Deep learning-based time series forecasting. *Artificial Intelligence Review.* 58. 2024. 10.1007/s10462-024-10989-8.

[20] Chen, Weisi & Hussain, Walayat & Cauteruccio, Francesco & Zhang, Xu. Deep Learning for Financial Time Series Prediction: A State-of-the-Art Review of Standalone and Hybrid Models. *Computer Modeling in Engineering and Sciences.* 139. 187-224. 2023.

10.32604/cmes.2023.031388.

[21] Vahid Naghashi, Mounir Boukadoum, and Abdoulaye Banire Diallo. A multiscale model for multivariate time series forecasting. *Scientific Reports*, 15:1565, 2025.

[22] Hongzhe Fang, Tianhong Li, and Huiting Xian. Comparative analysis of machine/deep learning models for single- step and multi- step forecasting in river water quality time series. *Water*, 17(13):1866, 2025.

[23] Suradhaniwar, Saurabh & Kar, Soumyashree & Durbha, Surya & Jagarlapudi, Adinarayana. Time Series Forecasting of Univariate Agrometeorological Data: A Comparative Performance Evaluation via One-Step and Multi-Step Ahead Forecasting Strategies. *Sensors.* 21. 2021. 10.3390/s21072430.

[24] Green, Riku & Stevens, Grant & Abdallah, Zahraa & Filho, Telmo. Stratify: unifying multi-step forecasting strategies. *Data Mining and Knowledge Discovery.* 39. 2025. 10.1007/s10618-025-01135-1.

[25] Eivind Strom and Odd Erik Gundersen. Performance metrics for multi- step forecasting measuring win-loss, seasonal variance and forecast stability: An empirical study. *Applied Intelligence*, 54:10490–10515, 2024.