DOCTORAL (PhD) DISSERTATION

LÁSZLÓ VANCSURA

Kaposvár

2025

HUNGARIAN UNIVERSITY OF
AGRICULTURE AND LIFE SCIENCES

HUNGARIAN UNIVERSITY OF
AGRICULTURE AND LIFE SCIENCES

THE ROLE OF MACHINE LEARNING MODELS IN FORECASTING ECONOMIC AND STOCK MARKET TIME SERIES

DOI: 10.54598/005800

DOCTORAL (PhD) DISSERTATION

LÁSZLÓ VANCSURA

Kaposvár
2025

Name of Doctoral School:	Doctoral School of Economic and Regional Sciences
Discipline:	Management and Business Administration
Head of Doctoral School:	Prof. Dr. habil Zoltán Bujdosó full professor Hungarian University of Agriculture and Life Sciences Institute of Rural Development and Sustainable Economy
Supervisor:	Dr. Tibor Bareith research fellow ELTE Centre for Economic and Regional Studies Institute of Economics

....

Approval of Head of Doctoral School Approval of Supervisor

TABLE OF CONTENTS

1. INTRODUCTION

Interest in the stock market and financial products among ordinary people has grown exponentially in recent decades (Kumbure et al., 2022). Billions of dollars worth of assets change hands every day on the world's stock exchanges (Hoseinzade and Haratizadeh, 2019), and investors enter the market with the intention of making a profit within their investment horizon. If a private or institutional investor could accurately predict market behavior and movements, this would enable them to achieve a higher risk-adjusted return (alpha) than the market. Among other things, this factor motivates the use of machine learning and artificial intelligence methods to create more accurate models for stock market forecasting and to fine-tune existing ones. The predictability of stock and other financial markets has been examined in numerous studies through the development of sophisticated forecasting systems (Sedighi et al., 2019; Song et al., 2019), some of which have reported that their models were able to generate significant profits (Atsalakis and Valavanis, 2009a; Weng et al., 2017).

In general, stock market forecasting is considered one of the most relevant yet challenging areas of financial research (Chen and Hao, 2017). Nevertheless, the ability of an investor to consistently achieve higher risk-adjusted returns than the market may violate the so-called efficient market hypothesis. Fama (1970) is credited with the market efficiency hypothesis (EMH). The EMH distinguishes between three forms of market efficiency. Weak form market efficiency assumes that information contained in past prices is already reflected in current stock prices and does not help predict future price movements (Fama, 1970). As a result, technical analysis cannot outperform a buy-and-hold strategy in terms of expected returns. The second form of the efficient market hypothesis is called semi-strong market efficiency, which states that stock prices reflect all publicly available information (economic conditions, political events, interest rates, company-specific information, etc.), including information about past prices. All this suggests that even using technical analysis tools, it is not possible to consistently achieve higher expected returns. In the case of semi-strong market efficiency, publicly available information, including fundamental data, does not allow an investor to outperform the market. This means that, with all publicly available information at their disposal, actively managed portfolios will not consistently achieve higher risk-adjusted returns than passive portfolios, i.e., those following a buy-and-hold strategy. The third, strong form of the EMH states that all information, including insider information, is reflected in stock prices. This precludes any investor (even one with insider information) from consistently achieving higher expected returns than the market (Fama, 1965). For this reason, the strong form of EMH essentially asserts that stock market prices and returns cannot be predicted (Timmermann and Granger, 2004). The strong form of EMH is based on extremely strict criteria, which Fama (1970) himself later partially refuted and refined. He stated that it cannot be expected that insider information cannot be used to realize higher expected profits.

Over time, more and more people have questioned the efficient market hypothesis and whether securities are priced rationally (Daniel et al., 1998; Borovkova and Tsiamas, 2019). There are numerous market anomalies that contradict the efficient market hypothesis (Malkiel and Mullainathan, 2005). These include financial market overreactions (De Bondt and Thaler, 1985) and underreactions, short-term momentum, long-term reversal, and high asset price volatility. Some researchers have discussed explanations for such anomalies that are consistent with the efficient market hypothesis, including that overreactions and underreactions occur randomly and with equal frequency (Fama, 1998). They have also examined the possibility that institutional investors (smart money) are able to offset the anomalies created by less experienced investors (dumb money) (Shiller, 2003). However, it remained doubtful whether models based on investor rationality could accommodate the observed anomalies. This led to a shift towards models that also integrate human psychology and to the emergence of behavioral finance, which questions the perfect rationality of investors due to behavioral biases such as loss aversion, overreaction, and underreaction. One attempt to reconcile EMH and behavioral finance was the proposal of the adaptive markets hypothesis (AMH), which acknowledges and explains the anomalies observed in financial markets (Lo, 2004).

Given that anomalies may exist, it is not surprising that most market participants use historical price information and company-specific information (past earnings, losses, and profits) as well as other factors to estimate future stock prices (Patel and Marwala, 2006). Stock market forecasting studies typically use two well-known analytical approaches: fundamental analysis and technical analysis (Lohrmann and Luukka, 2019; Sedighi et al., 2019). Fundamental analysis focuses on basic information. When forecasting a company's stock price or yield based on fundamentals, information such as the company's revenues and expenses, annual growth rate, market position, and other relevant factors included in financial statements or reports are taken into account. When forecasting a stock index that represents a number of company stocks, information about the market environment can also be used, including national economic production data, trade, prices, or interest rates that are likely to affect the performance of the companies included in the stock index. In contrast, technical analysis involves studying past trends in stock prices and related trading information (volume) in order to predict stock price movements (Wei et al., 2011). Based on the literature on the subject, it can be concluded that there are a number of models available for predicting prices, yields, and volatility, which researchers classify into three main groups. The first group includes traditional statistical methods, the second group includes methods based on some form of artificial intelligence, and the third group includes so-called hybrid methods (Kim & Won, 2018; Vidal & Kristjanpoller, 2020; Zolfaghari & Gholami, 2021).

The application of artificial intelligence (AI) and machine learning (ML) is fundamentally transforming the financial sector, which directly or indirectly affects other industries as well. Financial service providers are allocating significant investments to develop and improve data science-related areas. Since the 2007-2008 financial crisis, data-driven innovation and regulation have received particular attention, leading to a re-evaluation of banking and trading practices. Alternative data, such as voice recordings and social media posts, are playing an increasingly important role in decision-making, but analyzing such data poses a challenge for traditional approaches, which is why machine learning models have come to the fore. These algorithms provide the necessary computing power and flexibility to uncover complex patterns. Recent developments have enabled the effective application of scientific theories to make more accurate predictions.

Most previous studies have used some form of statistical time series method to predict stock market products, based on historical data (Efendi et al., 2018). Among these, autoregressive conditional heteroscedasticity (ARCH), autoregressive moving average (ARMA), and autoregressive integrated moving average (ARIMA) models, Kalman filtering, and exponential smoothing are the most popular techniques (Chen & Chen, 2015; Yeh et al., 2011). Later, with the advent of artificial intelligence (AI) and soft computing, these techniques received increasing attention in studies dealing with stock market forecasting. Unlike traditional time series methods, these techniques are capable of handling the nonlinear, chaotic, noisy, and complex data of the stock market, which can result in more accurate forecasts (Chen & Hao, 2017).

These methods represent innovative and advantageous alternatives, making them attractive to researchers for financial market forecasting. The shortcomings of the various methodologies have given rise to a third category, which includes so-called hybrid predictive models. These combine traditional statistical and machine learning-based methods to achieve the most accurate estimation results possible (Reston et al., 2014; Büyükşahin and Ertekin, 2019).

In my research, I will examine the most modern predictive models, which are an important tool for investor groups and companies in the areas of risk management, yield maximization, and profit maximization. For empirical analysis, I will use the most popular financial instruments, such as stock indices (S&P500, DAX, Nikkei225), commodity market products (crude oil, gold, silver), cryptocurrencies (Bitcoin, Ethereum, Litecoin), and currency pairs (EUR/USD, GBP/USD, AUD/USD) for the period from January 1, 2016, to June 30, 2022. In terms of testing the robustness of the models, it is important that this period includes the calm period (2018), Covid19 (2020) and the Russian-Ukrainian conflict (2022). Since cryptocurrencies are relatively new products compared to others, their price data also covers a shorter period, which is one of the reasons for choosing the start and end points of the database. For the modeling, I examine three deep learning algorithms (RNN, LSTM, GRU) and three hybrid methodologies created from them (LSTM-GRU, RNN-LSTM, RNN-GRU) using regression analysis. I evaluate the differences between actual and estimated prices using the MAPE indicator. I begin by presenting the results of the analyses by product type, followed by a comparison by product and model type. In the rest of this section, I compare single- and multi-variable methods, the absolute best and worst predictive performance, and the effects of activation function optimization. I will also use the predictive results of the models to develop a trading strategy, which I will compare with the buy-and-hold method. In this way, I will attempt to emphasize the practical economic usefulness of the thesis with quantifiable investment results.

2. OBJECTIVES

Nothing illustrates the spread of artificial intelligence better than the fact that various learning algorithms are slowly seeping into different areas of our lives, making our daily routines easier and processes more efficient. The advantages and disadvantages of this will, of course, generate a lot of debate, but I believe that the advance of these technologies is inevitable, especially in industries with high capital strength. Predicting the prices of various investment products has always been a challenge for both statisticians and financial professionals (Nabipour et al., 2020). The main goal of developing predictive models is to estimate market-generated uncertainties as accurately as possible, thereby minimizing risk factors.

The spread and increasingly widespread use of machine learning methodologies has contributed to improving the performance of predictive models and increasing the accuracy of forecasts (Maqsood et al., 2019). Experts involved in prediction face a number of fundamental challenges in model development. Issues such as complexity, noisy information, developmental characteristics, and non-linear relationships can be attributed to the instability of stock and financial markets, as well as the interrelationships between investor psychology and market behavior (Duarte et al., 2017).

In the development of predictive models, machine learning tools are therefore becoming increasingly important, helping investors, traders, and corporate risk managers to make optimal decisions. The primary goal of these methods is to learn and then automatically recognize different patterns in large data sets. The most advanced deep learning algorithms are constantly evolving, enabling them to predict price fluctuations more and more effectively in order to optimize various strategies.

The importance of risk management is particularly heightened during periods of high volatility, such as the 2008 global crisis, Covid19, or the stock market crash caused by the Russian-Ukrainian war. The unpredictability of the inflationary environment creates an additional need for the most effective tools possible. Today, the most modern risk management techniques go beyond traditional diversification, with artificial intelligence-based solutions becoming increasingly prominent and an integral part of our everyday lives. In the case of trading strategies, price forecasting models can determine key price levels that can be used in fundamental and technical analysis, as well as in risk management and portfolio management. The main goal of my research is to explore the characteristics of predictive modeling using machine learning models. I will describe the detailed objectives and hypotheses below.

C1: The primary objective of my research is to determine the extent to which different neural deep learning models can be generalized, i.e., whether they are capable of achieving outstanding predictive performance in different crisis situations.

For this reason, I will analyze three different periods, with 2018 representing a period of calm, while Covid19 (2020) and the outbreak of the Russian-Ukrainian conflict (2022) represent periods of crisis.

K1: What relationship can be demonstrated between the volatility of financial instruments and the predictive performance of price forecasting models?

K2: What kind of forecasting distortions are caused by the extreme price movements observed during periods of crisis, and how do different types of algorithms respond to this?

K3: What role does hybrid model architecture play in predictive performance change in a volatile market environment?

During the modeling process, I also discuss which of the product categories examined is the most stable and which poses the greatest challenge for predictive algorithms. There are significant differences between product specifics, as the analysis includes traditional stock markets with a long history, as well as cryptocurrencies, which are still in their infancy in economic terms but are all the more volatile. Commodity market products also play a role in the study, as their prices are influenced by a multitude of external factors. Foreign exchange markets should not be forgotten either, as they also have special characteristics, with prices being determined by a number of interrelated factors, such as interest rate differentials, inflation, political stability, and trade relations. Correlations between currency pairs can also have a significant impact on the accuracy of forecasts. Stock indices show more predictable trends in the longer term, while cryptocurrencies and certain commodity prices can be much more sensitive to short-term events, such as economic cycles or extreme events. This volatility poses a significant challenge for predictive models, as it is more difficult to make accurate forecasts in markets where prices can change rapidly and significantly. Stress factors affecting the global economy have a significant impact on the performance of various financial markets, so it is particularly important to be able to estimate the prices of the products under review as accurately as possible during such periods. One of the advantages of using machine learning models is that, compared to other methods, they are better at recognizing patterns in large amounts of historical data and using them to predict future prices and trends. A characteristic feature of deep learning models is that the more data they have at their disposal, the more effectively they can learn. It is therefore particularly important to teach them using data sets from economic periods with specific characteristics. I am seeking answers to the following research questions:

C2: The second objective of my thesis is to examine the extent to which machine learning models can improve the trading performance of financial instruments compared to the traditional buy-and-hold strategy. I seek answers to the following research questions:

K4: How well can trading strategies based on machine learning predictions exploit market anomalies, as opposed to the efficient market theory?

K5: What effect does the volatility of different asset classes (e.g., stocks, cryptocurrencies, commodities, currency pairs) have on the performance of machine learning-supported trading strategies?

K6: What differences can be observed in the performance of machine learning-based strategies during different economic cycles?

I use the forecasts to develop rule-based trading strategies. I backtest the performance of these strategies and compare them with the results of a passive investment approach, with a particular focus on returns and risk indicators (cumulative return, Sharpe ratio, etc.). The goal is not only to evaluate the accuracy of the predictions, but also to explore whether the decisions generated by the models actually result in improved trading performance. Through this, I seek to answer the question of how well machine learning-based decision support can handle different market conditions (trends, volatility, shocks). The practical value of the thesis is that the approach examined offers a potential alternative to classic passive investing. This is particularly relevant in the context of the efficient market hypothesis (EMH), which states that market prices reflect all available information, making it impossible to achieve extra returns in the long run using any predictive method. If the strategies presented in this paper are able to systematically outperform the buy-and-hold benchmark, they may implicitly call into question the practical validity of the EMH. In addition, maximizing the Sharpe ratio plays a central role in the evaluation of strategies, as this ratio measures the ratio of return to return volatility and is thus suitable for objectively comparing risk-weighted performance. The goal of the models is therefore not merely to achieve high returns, but to ensure stable, volatility-adjusted profitability. Examining this can help us understand the extent to which machine learning-based trading systems can create value in an environment where information efficiency theoretically precludes arbitrage opportunities.

3. MATERIALS AND METHODS

3.1. Data

In my research, I used daily exchange rate data for stock indices (S&P500, DAX, Nikkei225), commodity products (crude oil, gold, silver), cryptocurrencies (Bitcoin, Ethereum, Litecoin), and currency pairs (EUR/USD, GBP/USD, AUD/USD) for the period between January 1, 2016, and June 30, 2022. For crude oil, gold, and silver, I used futures prices, while for the other products, I used spot prices. I chose this period partly because it includes the calm period (2018), the Covid19 (2020) and the war crisis (2022), and partly because cryptocurrencies are relatively new products compared to the others, so their exchange rate data covers a shorter period. Therefore, this seemed to be the most optimal decision in terms of comparability. I collected the data from the website www.finance.yahoo.com, with the exception of cryptocurrencies, as their data comes from the website www.coinmarketcap.com.

During data cleaning and data series review, I also had to deal with the problem of handling missing data, which was particularly significant in the calculation of the correlation matrix, as the number of observations differed for each product. To solve this problem, I chose the linear interpolation method. In the case of forecasting methods, missing data was not significant, as each product was examined separately. I will discuss the filtering of trend and seasonal effects later when determining volatility. When making predictions, I did not treat trends and seasonality separately, but left this to the pattern recognition capabilities of the models. I then divided the databases into three parts. The first period examined focused on a calm economic environment from January 1, 2016, to June 30, 2018. The second covers the period from January 1, 2018, to June 30, 2020, which was selected due to

Covid19. The third period also relates to an economic crisis, namely the Russian-Ukrainian conflict, covering the interval between January 1, 2020, and June 30, 2022. During the analysis, the two crisis periods served as important factors in terms of testing the robustness of the models. The descriptive statistics of the data used for the empirical analysis are presented in Tables 1-3.

Table 1: Descriptive statistics of the examined products for the period between January 1, 2016 and June 30, 2018

	N	Average	Median	Std	Min	Max
S&P500	628	2360.37	2365.41	258.18	1829.08	2872.87
DAX	632	11575.03	12002.46	1243.42	8752.87	13559.60
Nikkei225	613	19315.52	19383.84	2355.25	14952.02	24124.15
Crude oil	626	50.81	49.56	9.54	26.21	74.15
Gold	625	1266.21	1271.50	57.64	1073.90	1364.90
Silver	625	16.98	16.87	1.25	13.74	20.67
Bitcoin	912	3649.33	1187.47	4189.45	364.33	19497.40
Ethereum	911	233.20	44.89	298.89	0.92	1398.99
Litecoin	912	52.79	6.93	72.31	3.00	358.34
EUR/USD	649	1.14	1.13	0.05	1.04	1.25
GBP/USD	649	1.33	1.32	0.07	1.20	1.48
AUD/USD	649	0.76	0.76	0.02	0.69	0.81

Source: own editing

In the first examined period (Table 1), the stock markets exhibited moderate growth and relatively stable volatility. The average value of the S&P 500 index was 2360.37 points, with a standard deviation of 258.18 points, whereas the DAX and Nikkei225 recorded higher averages and deviations, indicating greater price fluctuations in the European and Asian markets. The average price of crude oil was USD 50.81, accompanied by a relatively high standard deviation (USD 9.54), reflecting the commodity market’s sensitivity to geopolitical and supply–demand dynamics. Gold and silver functioned as more stable stores of value, with comparatively low volatility, particularly in the case of silver (USD 1.25). Cryptocurrencies, especially Bitcoin and Ethereum, displayed substantial volatility during this period: the standard deviation of Bitcoin was USD 4189.45, while that of Ethereum was USD 298.89. This elevated volatility underscores the rapid price surge and speculative nature of digital assets. In the foreign exchange markets, the EUR/USD, GBP/USD, and AUD/USD exchange rates demonstrated relative stability, with low deviations and narrow minimum–maximum ranges. Overall, this period was characterized by stability in traditional financial assets, while cryptocurrencies experienced pronounced growth and heightened volatility.

Table 2: Descriptive statistics of the examined products for the period between January 1, 2018 and June 30, 2020

	N	Average	Median	Std	Min	Max
S&P500	628	2862.46	2843.11	195.87	2237.40	3386.15
DAX	627	12099.54	12253.15	911.70	8441.71	13789.00
Nikkei225	606	21865.58	21874.23	1291.52	16552.83	24270.62
Crude oil	628	56.18	58.30	13.28	-37.63	76.41
Gold	627	1394.23	1328.10	159.39	1176.20	1793.00
Silver	627	16.06	16.14	1.29	11.73	19.39
Bitcoin	912	7679.92	7679.97	2384.93	3236.76	17527.00
Ethereum	911	302.16	207.80	238.30	81.72	1398.99
Litecoin	912	80.24	61.06	48.48	23.46	296.45
EUR/USD	651	1.14	1.13	0.04	1.07	1.25
GBP/USD	651	1.30	1.29	0.05	1.15	1.43
AUD/USD	651	0.71	0.71	0.04	0.57	0.81

Source: own editing

In the second period (Table 2), which partly overlaps with the onset of the Covid19 pandemic, increasing volatility and structural reconfiguration were observed in the financial markets. Stock indices, particularly the S&P 500 (average: 2862.46) and the DAX (average: 12099.54), exhibited moderate growth compared to the previous period; however, their standard deviations slightly declined, which may indicate partial market stabilization prior to the outbreak. In the crude oil market, an extraordinary price collapse occurred, as reflected in the negative minimum value (-37.63 USD) recorded in April 2020 on the futures market. Precious metals, especially gold, appreciated as safe-haven assets: the average price of gold rose to 1394.23 USD, while its standard deviation increased substantially (159.39 USD). Cryptocurrencies, such as Bitcoin and Ethereum, continued to display high volatility, although their rising median values were more moderate than before, suggesting certain signs of market maturation. Litecoin also registered an increase in its average price, accompanied by relatively high volatility. Among currencies, the EUR/USD and GBP/USD pairs showed slight depreciation, while the AUD/USD experienced a more pronounced decline, attributable to the Australian economy's dependence on commodity exports. Overall, this period reflected heightened uncertainty, but at the same time, adaptation in both traditional and novel financial asset markets.

Table 3: Descriptive statistics of the examined products for the period between January 1, 2020 and June 30, 2022

	N	Average	Median	Std	Min	Max
S&P500	628	3851.17	3915.98	599.47	2237.40	4796.56
DAX	635	13907.57	13950.04	1633.03	8441.71	16271.75
Nikkei225	606	26031.83	27007.45	3214.96	16552.83	30670.10
Crude oil	628	63.11	62.57	25.23	-37.63	123.70
Gold	628	1803.00	1809.30	104.27	1477.30	2051.50
Silver	628	22.99	23.97	3.70	11.73	29.40
Bitcoin	911	30776.76	33798.01	18183.70	4970.79	67566.83
Ethereum	911	1745.93	1790.25	1384.34	109.21	4800.00
Litecoin	911	116.99	109.43	68.26	30.93	386.45
EUR/USD	651	1.15	1.16	0.05	1.04	1.23
GBP/USD	651	1.32	1.33	0.06	1.15	1.42
AUD/USD	651	0.72	0.72	0.04	0.57	0.80

Source: own editing

The third period (Table 3) brought about profound market shifts driven by the Covid19 pandemic, the subsequent recovery, and the geopolitical uncertainty arising from the Russian–Ukrainian conflict. The average value of the S&P 500 index increased to 3851.17 points, while its standard deviation rose substantially to 599.47 points, indicating heightened market volatility. The DAX and Nikkei225 indices also recorded higher averages and maximum values, accompanied by rising volatility, particularly in the case of the Nikkei225. The crude oil market exhibited extreme sensitivity to the Russian–Ukrainian conflict: the standard deviation increased to 25.23 USD, and the maximum price reached 123.70 USD. The disruption and uncertainty surrounding Russian commodity exports, especially energy resources, played a key role in driving these fluctuations. Demand for precious metals, especially gold, intensified as investors sought safe-haven assets, although prices stabilized at elevated levels, with an average of 1803.00 USD. Bitcoin and Ethereum reached peak valuations during this period: Bitcoin averaged 30776.76 USD, with an exceptionally high standard deviation of 18183.70 USD. This may partly reflect both heightened risk aversion and speculative capital inflows amid the turbulence of traditional markets. The EUR/USD and GBP/USD exchange rates remained relatively stable; however, a slight depreciation was observed due to the European economy's direct exposure to the conflict. Overall, this period was characterized by intensifying geopolitical tensions, energy market shocks, and the predominance of investor risk aversion.

In the course of the analysis, both univariate and multivariate methods were applied. In the univariate case, daily closing prices formed the basis of prediction, whereby the price at a given point in time was estimated using the data of the preceding 50 time steps (sequence). In the multivariate analysis, a similar approach was adopted, with the difference that, in addition to daily closing prices, opening, highest, and lowest prices were also incorporated, again using 50-day sequences.

3.2. Methods

Simple Recurrent Neural Network (RNN)

RNN is a type of artificial neural network consisting of three main parts: input, hidden, and output layers. There are two main differences compared to traditional networks. One is that the nodes in the same hidden layer of an RNN are connected to each other, and the other is that the inputs to the hidden layer contain the outputs of the input layer at the current time as well as the outputs of the hidden layer stored at the previous time. This special structure allows for a better description of dynamic temporal behavior in a time series sequence. Thus, an RNN can use previously learned information to recognize the current pattern, enabling more efficient modeling (Bai et al. 2021).

Gated Recurrent Unit (GRU)

GRU is a type of recurrent neural network (RNN) that excels at predicting time series. It is similar to another neural network model we discussed (LSTM), but GRU has lower computational requirements, which can greatly improve learning efficiency. Its input and output structures are the same as those of a simple RNN. The internal structure of the GRU unit contains only two gates: the zt update gate and the rt reset gate. The zt update gate determines the previous memory value saved at the current time, and the rt reset gate determines how the new input information should be combined with the previous memory value. Unlike the LSTM algorithm, the zt update gate can simultaneously forget and select the contents of the memory, which improves computational performance and reduces the required running time (Xiao et al. 2022).

Long-Short Term Memory (LSTM)

LSTM is a type of recurrent neural network (RNN) that is often used in research related to sequential data. Long-term memory refers to learning weights, while short-term memory refers to the internal states of cells. LSTM was created to solve the vanishing gradient problem in RNNs, and its main change is the replacement of the middle layer of the RNN with a block (LSTM block). The main feature of LSTM is the possibility of long-term affiliation learning, which was impossible in RNNs. To predict data related to the next time point, the network weights need to be updated, which requires maintaining data from the initial time interval. An RNN could only learn a limited number of short-term affiliations; however, RNNs cannot learn long-term time series. The LSTM, however, can handle these adequately. The structure of the LSTM model contains a set of recurrent subnetworks, which are called memory blocks. Each block contains one or more autoregressive memory cells and three multiple units (input, output, and forget), which perform continuous writing, reading, and control of cell operation (Ortu et al. 2022).

LSTM-GRU hybrid

The hybrid LSTM/GRU model based on LSTM and GRU retains the advantages of both models, reduces overfitting, and thus enables high-accuracy predictions (Zhao et al. 2023). In this model, the first hidden layer is the LSTM. Each LSTM neuron collects data and generates a weighted value. The data is then transferred from the LSTM to the GRU layer, which is the second hidden layer. On the way from the LSTM layer to the GRU layer, weighted values are generated again. Similarly, the data is then transferred to the third hidden layer (dense layer). A weighted value is also generated from the GRU layer to the dense layer. The dense layer is a normal neural network layer that is used to produce the output. From the dense layer, the data then goes to the output neuron (Islam and Hossain 2021).

RNN-LSTM hybrid

In order to exploit the strengths of RNN and LSTM and eliminate their weaknesses, the use of the RNN-LSTM hybrid model significantly improves the predictability of time series (Faru et al. 2023). In this algorithm, the first hidden layer is the RNN, in which neurons collect information and a weighted value is generated. The information is then transferred from the RNN layer to the second hidden layer, the LSTM. Weighted values are generated again on the way from the RNN layer to the LSTM. The data is then transferred to the third hidden layer (dense layer). Weighted values are also generated from the LSTM layer to the dense layer.

RNN-GRU hybrid

Although there are quite a few hybrid algorithms in the literature, there's not much research on the RNN-GRU combo. This model is very similar to the previous one, so the first hidden layer is the RNN, in which neurons collect information and a weighted value is generated. The information is then transferred from the RNN layer to the second hidden layer, the GRU. Weighted values are generated again on the way from the RNN layer to the GRU. The data is then transferred to the third hidden layer (dense layer). Weighted values are also generated from the GRU layer to the dense layer.

3.3. Softwares

The Python programming language (version 3.9) was used for model development, which was justified for several reasons. On the one hand, Python has a broad and rapidly evolving ecosystem, particularly in the areas of data analysis (e.g., Pandas, Numpy) and machine learning (e.g., TensorFlow, Keras, Scikit-learn, Pytorch), which enables the efficient implementation of state-of-the-art neural network models. On the other hand, Python is open source and platform-independent, ensuring cost-effective development and reproducibility. Another advantage is that the programming language's simple, easy-to-read syntax allows for rapid prototype development, which is particularly useful in research projects where models need to be continuously fine-tuned. Other advantages of Python include a wide range of frameworks, support for deep learning, and industrial applicability. I ran the scripts on the Google Colab interface and locally using the Visual Studio Code code editor.

3.4. Evaluation of models

To evaluate the models used for forecasting and determine their accuracy, the following metrics are most commonly used in the literature: root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE). In my dissertation, I used the MAPE indicator to measure deviations. This indicator measures the average magnitude of forecast errors and shows the deviations in percentage form.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

The lower the values of the above indicators, the more reliable and accurate the forecasts will be. MAPE is interpreted as a percentage (the deviations are expressed as a percentage of the original value). For this reason, MAPE can also be used to compare different instruments, as it does not depend on the nominal value of the exchange rate. Since our study examined indices from many parts of the world and the effects of two different negative economic events, we used the MAPE indicator for the comprehensive evaluation of the models for the sake of comparability.

3.5. Volatility analysis

The variance data for different products and periods are not suitable for comparing volatility and MAPE values, so it was necessary to use an indicator measuring relative volatility, the calculation of which is detailed below. First, it was necessary to find a method that identifies trends in time series and can remove them, as they can distort the results. For this, I used the Seasonal-Trend decomposition using Loess (STL) approach. I will discuss the methodology in detail below.

3.5.1. STL (Seasonal-Trend decomposition using Loess)

STL is a flexible time series analysis method that breaks down the series under examination into three components: the long-term trend, the seasonal pattern, and the random residual. A special feature of STL is the use of LOESS (Locally Weighted Scatterplot Smoothing), which is also capable of modeling nonlinear, time-varying seasonal patterns. Another advantage is its robustness and flexibility, making it well suited for analyzing financial time series, for example. The time series is decomposed in the following form:

$$Y_t = T_t + S_t + R_t$$

where Y_t is the original time series, T_t is the trend, S_t is the seasonal component, and R_t is the residual. STL can be effectively applied to financial time series where seasonal patterns may change over time (Cleveland et al., 1990). The residual R_t contains short-term fluctuations independent of the trend, making this component suitable for volatility analysis (He et al., 2021).

3.5.2. Relative volatility: coefficient of variation (CV)

I measured the volatility of the trend-adjusted series using the coefficient of variation (CV). The CV is a dimensionless indicator that allows the comparison of the dispersion of assets with different units of measurement and magnitudes. Its formula is as follows:

$$CV = \frac{\sigma(R_t)}{\mu(Y_t)}$$

where $\sigma(R_t)$ is the standard deviation of the residual component, while $\mu(Y_t)$ is the mean of the original time series. The indicator is sensitive to the size of the series average, so CV may show greater volatility for assets with low nominal values. CV strikes a balance between absolute and relative variance-based volatility measures and is well suited as a basis for cluster analysis (Brockwell and Davis, 2002).

Relative volatility is calculated from the residual components of the aforementioned method (STL). In the case of STL, we obtain trend- and seasonality-free residuals.

3.6. Prediction-based, dynamically optimized trading strategy

The algorithm I use is a parameter-sensitive trading strategy based on machine learning predictions, which is compared to the classic buy-and-hold approach. The central element of the strategy is that it generates trading positions based on the relationship between predicted and actual exchange rates, then optimizes the risk/return ratio using various stop-loss (SL), take-profit (TP), and volatility estimation window (VOL_LOOKBACK) values. Its goal is therefore to maximize trading performance according to the Sharpe ratio. The input for the source program is provided by a multi-page Excel file, which contains the following for each instrument: timestamp ("Date"), actual market values (closing prices), and estimated values generated by the prediction models. The strategy is based on the assumption that if the estimated value is higher than the actual value at a given point in time, the price is expected to rise, so it is worth taking a long position. Otherwise, a short position is justified. This can be formally described as follows: if $\hat{P}_t > P_{t-1}$, then opens a long position, as long as $\hat{P}_t < P_{t-1}$n

Log return is calculated as follows: $r_t = ln\left(\frac{P_t}{P_{t-1}}\right)$

The strategy is further refined by volatility-based risk management, for which three parameters are set:

- SL_MULTIPLIER: stop-loss threshold ratio compared to the volatility of the previous period.
- TP_MULTIPLIER: take-profit threshold ratio.
- VOL_LOOKBACK: the period length used to calculate volatility (moving standard deviation calculated on a specific day).

Volatility and limits are calculated as follows:

$$\sigma_t = std_{rolling}(r_t, VOL\_LOOKBACK)$$

$$SL_t = SL\_MULTIPLIER * \sigma_t$$

$$TP_t = TP\_MULTIPLIER * \sigma_t$$

The realized return will be adjusted based on these if the profit or loss reaches the specified levels. For example, if the return in a long position exceeds the TP, we will cut the return at the TP value; if the loss falls below the SL, we will cut it at the -SL value. The same applies to short positions.

Building on this base strategy, I incorporated and tested two other methods. In one case, it calculates a 5-day rolling MAPE indicator and only opens a position if the deviation indicator is below a specified threshold. In this case, I set the threshold at 3%, so the MAPE acts as a kind of filter for opening positions. The other alternative also uses the rolling MAPE indicator value, but only to determine the dynamic position size. Specifically, it takes a smaller position size in the case of a higher MAPE, while in the opposite case, it implements a larger allocation.

The program performs static and cumulative performance analysis, calculating the most important statistical indicators for each parameter combination and instrument examined. These include the average daily return (μ), the standard deviation of the daily return (σ), and the Sharpe ratio, which is calculated using the standard annualized formula: $\frac{\mu}{\sigma} * \sqrt{252}$. In addition, the function calculates the cumulative return, the maximum drawdown (which is the largest difference between the local maxima and minima of the cumulative return curve), and the win rate. The calculations are performed separately for the buy-and-hold benchmark strategy and the dynamically managed strategy, allowing for a multidimensional comparison of the two approaches. During the optimization process, the algorithm iterates through a predefined parameter grid. This grid contains the following values: SL $\in \{0.5, 1.0, 1.5\}$, TP $\in \{1.5, 2.0, 3.0\}$, and VOL_LOOKBACK $\in \{10, 20, 30\}$. The strategy is run for all possible combinations - 27 in total - and then the system selects the parameter configuration with the highest Sharpe ratio. The setting determined in this way is stored as a parameter set optimized for the given instrument. At the end of the analysis, the algorithm performs a graphical comparison in the form of a bar chart. This figure shows the cumulative returns of the four strategies for each instrument. This makes it clear whether the model is able to consistently outperform the passive investment strategy, i.e., whether the parameter-tuned trading approach based on the predictive model has practical value.

In summary, the program is a generally applicable tool for analyzing prediction-based investment strategies. Its main features include the explicit integration of the predictive component, the application of parametric risk management, automatic optimization, and systematic comparison with the classic buy-and-hold benchmark strategy. The code can be used not only in a simulation environment, but can also be effectively adapted for backtesting systems, machine learning model validation, and sensitivity testing.

When applying trading strategies, similar to most other literature, I did not take taxes, transaction and other costs into account, so the results represent gross cumulative returns.

4. RESULTS

4.1. Results of volatility analysis

The examination of the relationship between volatility and MAPE was conducted following the activation function optimization process. Based on volatility values, cryptocurrencies clearly stand out, particularly in 2018: Ethereum (0.0475) and Litecoin (0.0432) exhibit extremely high values, in contrast to the low CV levels characteristic of currency pairs (EUR/USD: 0.0032). In 2020, crude oil displayed exceptionally high volatility (0.1005), which can be attributed to the market shocks associated with the Covid19 period. Classical stock indices such as the S&P 500 and the DAX also showed higher volatility in 2020 compared to 2018 or 2022, indicating the intensification of market uncertainty.

According to the pooled regression analysis covering all periods (Figure 1), a significant positive relationship was identified between relative volatility (CV_STL) and univariate MAPE. The regression coefficient was 1.2937, implying that a one-unit increase in volatility increases the forecasting error by an average of 1.2937 units. The statistical significance of the coefficient ($p < 0.01$) confirms the robustness of this relationship. The constant term was 0.0049 but proved non-significant, suggesting that the level of MAPE is essentially explained by the level of volatility. This finding indicates that the forecasting performance of univariate models is substantially affected by market fluctuations. The explanatory power of the model was high, with an R^2 value of 76.28%. In terms of univariate MAPE values, the weakest predictive performance in 2018 was observed for Litecoin (0.1095), followed by Bitcoin (0.1011). In the same year, the lowest error values were associated with currency pairs, such as GBP/USD (0.0047) and AUD/USD (0.0055). In 2020, crude oil also stood out with a MAPE value of 0.1023, reinforcing the notion that turbulent market movements significantly impair forecasting performance.

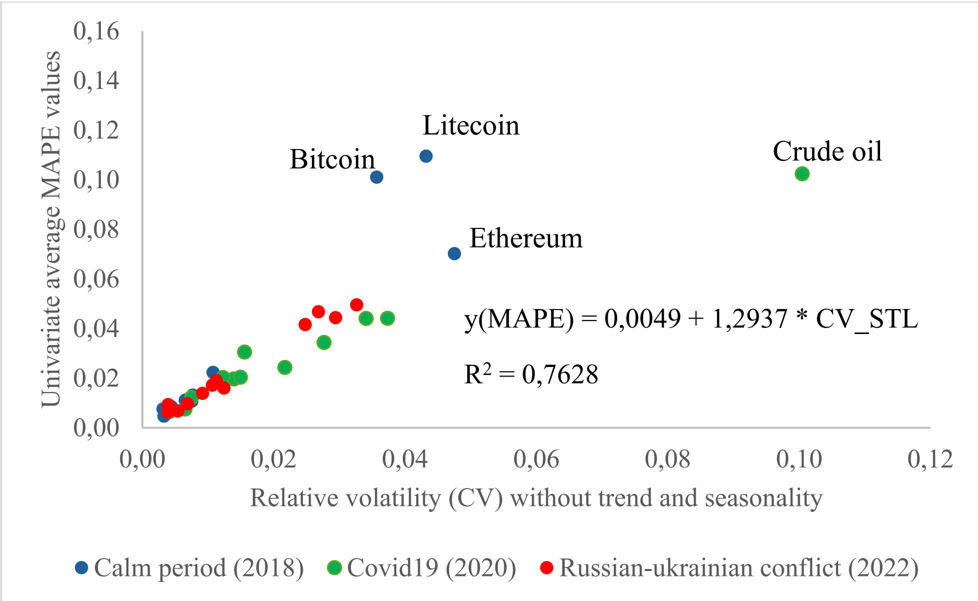

Figure 1: The relationship between relative volatility and univariate average MAPE values, displayed by time periods

Source: own editing

In the multivariate model (Figure 2), the effect of volatility appears even more pronounced: the regression coefficient of CV_STL is 1.6224, which is likewise significant ($p < 0.01$). This suggests that multivariate forecasting systems are even more sensitive to volatility, meaning that input complexity does not reduce, but in some cases actually increases error sensitivity under market fluctuations. The constant term is -0.0009 and again non-significant, indicating that relative volatility remains the decisive factor in explaining the variability of the model. The high coefficient value reinforces the conclusion that volatility continues to play a key role in shaping predictive performance, even in more advanced multivariate models. The explanatory power of the model was even higher than in the univariate case, with an R^2 of 90.58%. In terms of MAPE values, the trend is similar, but the errors are generally lower or nearly identical compared to the univariate models. Interestingly, in 2018 Litecoin's multivariate MAPE (0.1198) was even higher than its univariate value, which may point to issues related to predictor selection. In most cases, however, multivariate models slightly improve forecasting accuracy, as observed for the Nikkei225 in 2020 (0.0203 vs. 0.0204) and for GBP/USD in 2022 (0.0040 vs. 0.0063).

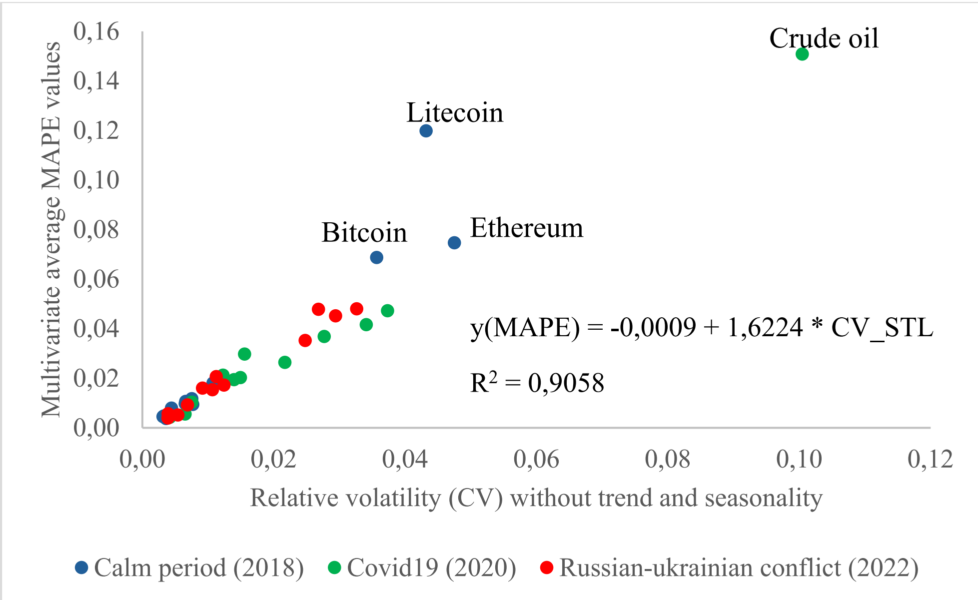

Figure 2: The relationship between relative volatility and multivariate average MAPE values, displayed by time periods

Source: own editing

The relationship between relative volatility and MAPE values, differentiated by model and broken down for the periods 2018, 2020, and 2022, is presented in Tables 4–6, based on the STL approach. These results are described in detail below.

Table 4 shows the relationship between relative volatility (CV_STL) and both univariate and multivariate forecasting errors (MAPE) for the year 2018, across six neural network models. The results indicate that in all univariate models, a significant positive relationship exists between volatility and MAPE, meaning that the higher the volatility of a given financial asset, the greater the prediction error. Among the univariate analyses, the highest regression coefficient was observed in the LSTM model (2.920), while the lowest was recorded for the GRU model (1.528), reflecting the differing sensitivities of network architectures to volatility. Based on R^2 values, the GRU model achieved the strongest explanatory power ($R^2 = 0.9713$), whereas the RNN showed the weakest ($R^2 = 0.6134$). These findings confirm that relative volatility is a strong and stable predictor of univariate forecasting performance. The second part of the table presents the relationship between relative volatility and multivariate forecasting errors (MAPE). The results suggest that in all models there is again a significant positive relationship between volatility and forecasting error, indicating that more volatile assets are associated with higher prediction errors, even when additional input variables are included. The highest regression coefficient was found in the RNN-LSTM model (2.856), while the lowest was observed in the LSTM-GRU model (1.662), highlighting that certain architectures are more sensitive to the effects of volatility. In terms of R^2 values, the LSTM-GRU model demonstrated the best fit (0.9651), while the RNN-LSTM showed the weakest (0.7951). Overall, it can be concluded that both univariate and multivariate models display a strong and consistent relationship between relative volatility and prediction error.

Table 4: Regression results of the relative volatility indicators and MAPE values for the examined assets during the calm period (2018)

	(1) Univ-RNN	(2) Univ-LSTM	(3) Univ-GRU	(4) Univ-LSTM-GRU	(5) Univ-RNN-LSTM	(6) Univ-RNN-GRU
CV_STL	2.422***	2.920***	1.528***	2.128***	2.304***	1.817***
	(0.608)	(0.578)	(0.083)	(0.315)	(0.168)	(0.275)
Constant	-0.002	-0.006	-0.000	-0.001	-0.003	0.002
	(0.013)	(0.013)	(0.002)	(0.007)	(0.004)	(0.006)
Observations	12	12	12	12	12	12
R^2	0.6134	0.7185	0.9713	0.8207	0.9498	0.8132
	(7) Multi-RNN	**(8) Multi-LSTM**	**(9) Multi-GRU**	**(10) Multi-LSTM-GRU**	**(11) Multi-RNN-LSTM**	**(12) Multi-RNN-GRU**
CV_STL	1.830***	2.854***	1.810***	1.662***	2.856***	1.870***
	(0.264)	(0.346)	(0.188)	(0.100)	(0.458)	(0.239)
Constant	0.000	-0.008	-0.004	-0.001	-0.007	-0.001
	(0.006)	(0.008)	(0.004)	(0.002)	(0.010)	(0.005)
Observations	12	12	12	12	12	12
R^2	0.8273	0.8720	0.9022	0.9651	0.7951	0.8595

Source: Own editing based on STATA 17 results

In parentheses, the standard errors

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5 presents the relationship between relative volatility (CV_STL) and both univariate and multivariate forecasting errors (MAPE) for the year 2020, across six neural network models. For all univariate models, the regression coefficient of volatility is significant and positive, confirming that higher volatility is associated with increased prediction error during this crisis-laden period. The lowest coefficient was observed in the GRU model (0.845), while the highest was recorded for the RNN model (1.158), reflecting the differing sensitivities of the architectures to volatility. The R^2 values are extremely high across all models (each exceeding 90%), indicating outstanding explanatory power of the regression analysis during the period under review. The second part of the table illustrates the relationship between relative volatility (CV_STL) and multivariate forecasting errors (MAPE). Here too, the results indicate a statistically significant and positive relationship across all models, meaning that prediction errors in multivariate forecasts also increase with rising volatility. The highest regression coefficient was observed for the RNN-LSTM model (2.847), while the lowest was found for the GRU (0.828), which may suggest greater robustness of the latter to volatility. The coefficients of determination (R^2) are remarkably high for all models (ranging between 86% and 98%), with the LSTM model in particular (0.9714) exhibiting the strongest explanatory power. These results confirm that even in the pandemic-stricken year of 2020, a strong relationship persisted between the extent of market fluctuations and the forecasting errors of the models, including the more advanced hybrid architectures.

Table 5: Regression results of the relative volatility indicators and MAPE values for the examined assets during the Covid19 period (2020)

	(1) Univ-RNN	**(2) Univ-LSTM**	**(3) Univ-GRU**	**(4) Univ-LSTM-GRU**	**(5) Univ-RNN-LSTM**	**(6) Univ-RNN-GRU**
CV_STL	1.158^{***} (0.067)	0.964^{***} (0.044)	0.845^{***} (0.077)	0.984^{***} (0.046)	0.984^{***} (0.100)	0.936^{***} (0.064)
Constant	0.005^{*} (0.002)	0.004^{**} (0.002)	0.006^{**} (0.003)	0.005^{***} (0.002)	0.011^{**} (0.004)	0.007^{***} (0.002)
Observations	12	12	12	12	12	12
R^2	0.9676	0.9795	0.9228	0.9787	0.9067	0.9549
	(7) Multi-RNN	**(8) Multi-LSTM**	**(9) Multi-GRU**	**(10) Multi-LSTM-GRU**	**(11) Multi-RNN-LSTM**	**(12) Multi-RNN-GRU**
CV_STL	2.279^{***} (0.158)	0.896^{***} (0.049)	0.828^{***} (0.066)	0.911^{***} (0.050)	2.847^{***} (0.289)	1.150^{***} (0.146)
Constant	-0.010^{*} (0.006)	0.004^{**} (0.002)	0.005^{*} (0.002)	0.004^{**} (0.002)	-0.020^{*} (0.010)	0.005 (0.005)
Observations	12	12	12	12	12	12
R^2	0.9543	0.9714	0.9403	0.9706	0.9065	0.8616

Source: Own editing based on STATA 17 results
In parentheses, the standard errors
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6 presents the relationship between relative volatility (CV_STL) and both univariate and multivariate forecasting errors (MAPE) for the year 2022. In all univariate models, a positive and significant relationship is observed, meaning that increases in volatility continue to be associated with higher forecasting errors. The regression coefficients range from 1.149 (GRU) to 1.969 (RNN-LSTM), indicating a relatively consistent, moderate sensitivity to volatility. The explanatory power (R^2) is high across all models, ranging between 85% and 92%. The second part of the table represents the relationship between relative volatility (CV_STL) and multivariate forecasting errors (MAPE). Here as well, all models exhibit significant and positive regression coefficients. The coefficients fall within a relatively narrow range, with the lowest value observed for the GRU model (1.312) and the highest for the RNN-LSTM (2.033), reflecting the differing sensitivities of network architectures to volatility. The R^2 values are high, ranging between 82% and 97%, with particularly strong explanatory power for the LSTM-GRU model (97.48%). The results consistently reinforce the findings of previous years: the structural presence of volatility fundamentally shapes forecasting performance, even in more complex multivariate deep learning models.

Table 6: Regression results of the relative volatility indicators and univariate MAPE values for the examined assets during the Russian–Ukrainian conflict period (2022)

	(1) Univ-RNN	(2) Univ-LSTM	(3) Univ-GRU	(4) Univ-LSTM-GRU	(5) Univ-RNN-LSTM	(6) Univ-RNN-GRU
CV_STL	1.856^{***} (0.205)	1.403^{***} (0.132)	1.149^{***} (0.133)	1.539^{***} (0.192)	1.969^{***} (0.254)	1.598^{***} (0.165)
Constant	-0.002 (0.004)	-0.000 (0.002)	0.003 (0.002)	0.000 (0.003)	-0.001 (0.005)	-0.001 (0.003)
Observations	12	12	12	12	12	12
R^2	0.8915	0.9192	0.8819	0.8656	0.8570	0.9032
	(7) Multi-RNN	**(8) Multi-LSTM**	**(9) Multi-GRU**	**(10) Multi-LSTM-GRU**	**(11) Multi-RNN-LSTM**	**(12) Multi-RNN-GRU**
CV_STL	1.370^{***} (0.153)	1.523^{***} (0.225)	1.312^{***} (0.115)	1.395^{***} (0.071)	2.033^{***} (0.225)	1.903^{***} (0.188)
Constant	0.003 (0.003)	-0.000 (0.004)	-0.001 (0.002)	-0.001 (0.001)	-0.003 (0.004)	-0.003 (0.003)
Observations	12	12	12	12	12	12
R^2	0.8889	0.8206	0.9283	0.9748	0.8904	0.9112

Source: Own editing based on STATA 17 results
In parentheses, the standard errors
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

To enable a comprehensive quantifiable assessment and comparison of the models, I also conducted a panel regression analysis for both univariate and multivariate approaches across all examined periods. The results are presented in Table 7. In exploring the relationship between univariate MAPE and relative volatility (CV_STL), I found that all coefficients are positive and statistically significant ($p < 0.01$), indicating that rising volatility systematically corresponds with increased prediction errors in the panel regression model as well. The highest sensitivity was observed for the RNN model (1.485), while the lowest was found for the GRU model (0.923). The R^2 values reflect varying explanatory power: the GRU model achieved the best fit (0.8632), while the LSTM yielded the weakest (0.5259), highlighting the heterogeneous performance of the models. For the multivariate neural network models, a positive and significant relationship between relative volatility and MAPE values was also observed across all aggregated panel regressions. The highest regression coefficient was recorded for the RNN-LSTM model (2.719), while the lowest appeared in the GRU (1.035), once again pointing to the somewhat more robust behavior of GRU models under higher volatility conditions. The R^2 values ranged between 60% and 90% depending on the model, thus showing variable but generally strong explanatory power, particularly in the case of the RNN model (90.72%), which exhibited outstanding fit. Overall, the results demonstrate that both univariate and multivariate deep learning models are systematically influenced by volatility in their forecasting errors, with this effect remaining consistent across different time periods. This reinforces the central role of volatility in shaping predictive performance, even within more complex model structures.

Table 7: Panel regression results of the relative volatility indicators and MAPE values for the examined assets across all periods

	(1) Univ-RNN	(2) Univ-LSTM	(3) Univ-GRU	(4) Univ-LSTM-GRU	(5) Univ-RNN-LSTM	(6) Univ-RNN-GRU
CV_STL	1.485***	1.418***	0.923***	1.274***	1.373***	1.185***
	(0.201)	(0.231)	(0.067)	(0.135)	(0.131)	(0.114)
Constant	0.004	0.003	0.006***	0.005	0.007*	0.006**
	(0.005)	(0.006)	(0.002)	(0.004)	(0.003)	(0.003)
Observations	36	36	36	36	36	36
R^2	0.6157	0.5259	0.8632	0.7251	0.7645	0.7606
	(7) Multi-RNN	**(8) Multi-LSTM**	**(9) Multi-GRU**	**(10) Multi-LSTM-GRU**	**(11) Multi-RNN-LSTM**	**(12) Multi-RNN-GRU**
CV_STL	2.144***	1.369***	1.035***	1.117***	2.719***	1.198***
	(0.118)	(0.192)	(0.104)	(0.073)	(0.188)	(0.112)
Constant	-0.007**	0.003	0.003	0.003	-0.012**	0.006*
	(0.003)	(0.005)	(0.003)	(0.002)	(0.005)	(0.004)
Observations	36	36	36	36	36	36
R^2	0.9072	0.5982	0.7794	0.8734	0.8604	0.8102

Source: Own editing based on STATA 17 results
In parentheses, the standard errors
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The aggregated panel regression results presented in Table 7 provide an opportunity to compare the prediction sensitivity of the models as a function of relative volatility (CV_STL). The positive regression coefficients across all models confirm that increases in volatility lead to higher forecasting errors (MAPE). At the same time, the magnitude of the coefficients allows us to infer which models are the least sensitive to volatility, i.e., more robust. Based on these findings, a comparison between univariate and multivariate models follows.

The results clearly show that in most model pairs, the multivariate versions respond more sensitively to volatility, as indicated by the higher CV_STL coefficients. For the RNN, GRU, RNN-LSTM, and RNN-GRU models, the univariate versions exhibit more favorable (lower) coefficients. However, for the LSTM and LSTM-GRU models, the opposite pattern is observed. Overall, in most cases, multivariate models demonstrate greater sensitivity to volatility, as reflected by higher regression coefficients. This suggests that while multivariate models generally achieve higher predictive performance (as also indicated by higher R^2 values, e.g., Multi-RNN: 0.9072 vs. Univ-RNN: 0.6157), they may be more vulnerable during periods of market turbulence, meaning that forecasting errors are more dependent on the level of volatility. Therefore, explicit handling of volatility dependence may be particularly warranted when applying these models.

Comparisons between simple and hybrid models also yield important insights. In the univariate analyses, based on results for the RNN and its hybrid variants, it can be concluded that the hybrid versions were less sensitive to volatility than the base model. The same pattern applies to the LSTM and its hybrid variants. In contrast, for the GRU, the opposite was observed: the conventional base type exhibited lower volatility sensitivity than the hybrid variants. In the multivariate analyses, the GRU base model also outperformed its hybrid variants. Comparisons of RNN and LSTM models are more mixed, as in some cases the hybrid models (RNN-GRU and LSTM-GRU) displayed lower sensitivity than their base counterparts. Based on the regression coefficients, it can be concluded that univariate models are, on average, less sensitive to volatility and therefore demonstrate more robust performance. Simple architectures (particularly GRU) exhibit lower volatility sensitivity than hybrid models. The performance of hybrid models spans a wide spectrum: while some (e.g., RNN-LSTM) are highly sensitive to volatility, others (e.g., RNN-GRU) are surprisingly stable. Consequently, model selection strongly influences how well a given architecture can manage the effects of volatility. This comparison highlights that not only the model type but also the input structure (univariate vs. multivariate) must be carefully considered when designing predictive systems, particularly in highly volatile market environments.

4.2. Comparison of trading strategy results

The basis of the trading strategies was formed by the predictive values generated by the models with the best forecasting performance for each period and asset category. The following section compares the returns and risk metrics of strategies based on backtesting actual versus predicted values.

Based on the 2018 data (Figure 3), it is evident that the traditional buy-and-hold strategy was outperformed in many cases by the more advanced machine learning-based models. This was particularly true for volatile asset classes such as cryptocurrencies and, to some extent, precious metals, although the predictive models also demonstrated advantages in equity and currency markets. For stock indices (S&P 500, DAX), the superiority of the MAPE-based models was clearly observable. For example, the buy-and-hold strategy for the S&P 500 achieved only a 0.83% cumulative return, whereas MAPE-based strategies yielded 12–14% returns with an outstanding Sharpe ratio around 2.65. In the case of the DAX, the buy-and-hold strategy produced a negative return (-17.12%), in contrast to the positive performance of the MAPE-based models, which not only improved returns but also exhibited a significantly more favorable risk profile. For the Nikkei225 index, however, all strategies generated negative returns. For gold and silver, a clear increase in returns was also observed. For instance, the MAPE-position strategy yielded a 7.02% return and a Sharpe ratio of 3.1859 for gold, while the buy-and-hold approach showed only marginal positive performance. The only exception was crude oil: the buy-and-hold strategy achieved a 20.56% positive return, whereas active strategies could not meaningfully improve on this. The predictive models' Sharpe ratios were negative, and maximum drawdowns reached considerably higher levels (e.g., MAPE filter: 0.5712), indicating that the model was unable to adequately track the nonlinear and turbulent movements of the oil market. The most pronounced differences were observed in cryptocurrencies. Bitcoin's buy-and-hold return was strongly negative (-73.38%), whereas the Base Strategy achieved over 86% cumulative return. Similarly significant improvements were noted for Ethereum and Litecoin, where machine learning-based strategies provided substantial return advantages, with Sharpe ratios consistently exceeding two. For currency pairs (EUR/USD, GBP/USD, AUD/USD), the machine learning models generally exhibited positive Sharpe ratios that were higher than those of the buy-and-hold strategies, even when absolute return levels remained moderate. Additionally, in nearly all cases, they resulted in lower maximum drawdowns. Overall, the Base, MAPE-filter, and MAPE-position strategies in 2018 outperformed the buy-and-hold approach across almost all asset classes, particularly in terms of returns and risk-adjusted performance (Sharpe ratio). Maximum drawdowns were also lower under the predictive models, reflecting more effective risk management. The exception of the oil market highlights that, while machine learning represents a promising tool for financial forecasting, there are market environments in which predictive performance is not guaranteed.

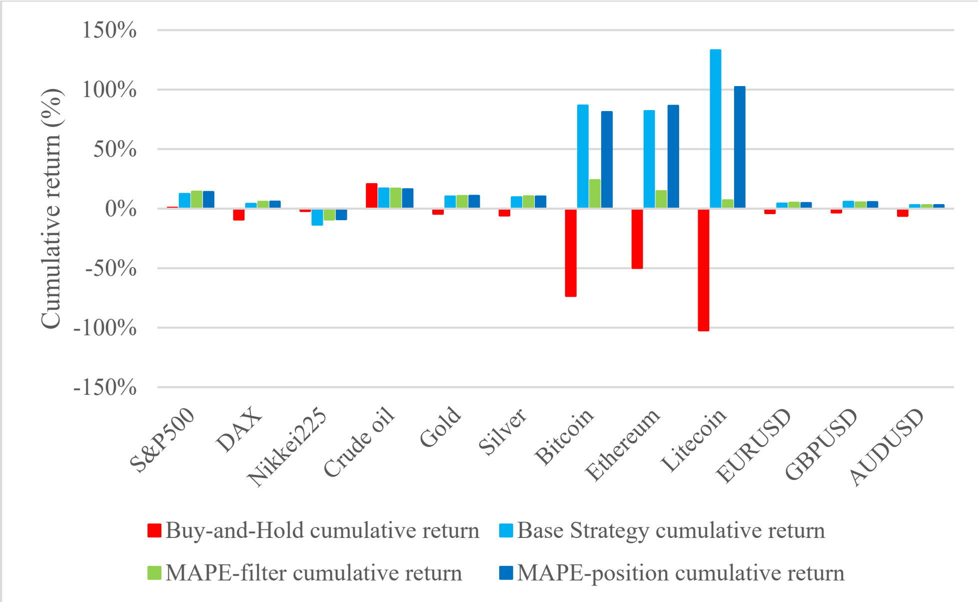

Figure 3: Comparison of the 4 trading strategies for the calm period (2018)

Source: own editing

The extreme market environment of 2020, marked by Covid19 (Figure 4), sharply highlighted the performance differences between various investment strategies. The traditional buy-and-hold approach underperformed in many asset classes compared to machine learning-based adaptive strategies, which not only provided higher returns but also better risk-adjusted performance and lower maximum drawdowns. The S&P 500 is a particularly clear example of the advantages of predictive strategies. While the buy-and-hold return was slightly negative (-5.65%), the Base Strategy achieved a cumulative return of 74.26%. The Sharpe ratio for the predictive models reached an extremely high value (4.7), compared to the negative ratio of buy-and-hold (-0.2465). Similar results were observed for the DAX and Nikkei225: alongside losses from the traditional strategy, the MAPE-based models delivered positive returns and significantly better risk metrics. Active strategies also performed well for precious metals. For gold, the MAPE-position strategy achieved a 23% return and a Sharpe ratio of 2.48, clearly outperforming buy-and-hold. Maximum drawdowns were also more moderate for the active models, reflecting improved risk management. The crude oil market again behaved as an exception in 2020. While the buy-and-hold strategy yielded a 17.12% return, the active strategies suffered substantial losses (-45.52%) and produced negative Sharpe ratios. This is likely attributable to market extremes (negative futures prices) that represented structural changes for which the predictive algorithms were not prepared. Following usual volatility patterns, the models often favored incorrect directions, which visibly worsened performance. In cryptocurrency markets, active approaches almost universally led to significant outperformance. Litecoin's buy-and-hold return was 1.23%, whereas the predictive strategy produced a cumulative return of 127.96% and a Sharpe ratio above 2.75. Similar trends were observed for Bitcoin and Ethereum. Machine learning-based strategies (Base Strategy and MAPE-position) achieved returns exceeding 100% with consistently strong Sharpe ratios between 2 and 3. For currency markets, the advantage of predictive strategies was smaller but still noticeable. Although absolute return differences were not dramatic, machine learning models generally improved risk-adjusted performance and reduced drawdowns in most cases. In 2020, amid extreme market volatility, machine learning strategies were not merely an alternative but, in many cases, a more effective and stable solution for portfolio management. Active strategies achieved outstanding results in equity markets, cryptocurrency markets, and precious metals, while producing moderate but stable improvements in currency markets. The oil market, however, again posed significant challenges to predictive systems.

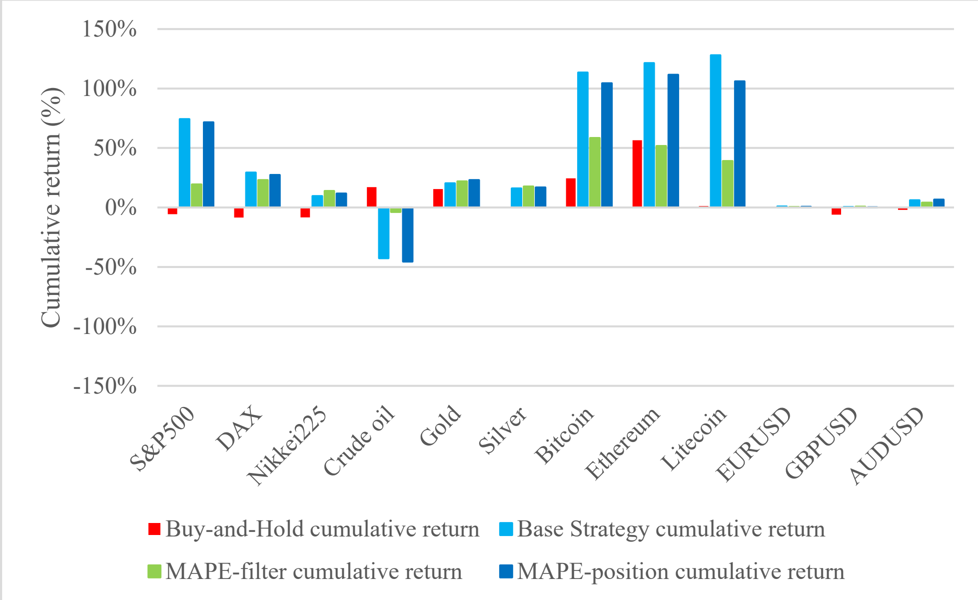

Figure 4: Comparison of the 4 trading strategies for the Covid19 period (2020)

Source: own editing

The market environment in 2022 (Figure 5) presented significant challenges for investors. Geopolitical tensions, inflationary pressures, and monetary tightening generated substantial volatility across traditional asset classes. In this uncertain context, predictive, machine learning-based strategies often provided effective protection against losses, and in many cases even generated positive returns where buy-and-hold suffered significant drawdowns. In equity markets, the buy-and-hold approach proved ineffective: the S&P 500 posted a -22.42% return with a Sharpe ratio of -1.81, whereas the Base Strategy achieved a positive return of 13.49% and a Sharpe ratio of 1.46. The contrast was even more pronounced for the DAX: buy-and-hold closed with -31.16% return and a -2.06 Sharpe ratio, while predictive strategies delivered positive, albeit more modest, performance. The superiority of active strategies was also evident for the Nikkei225, with buy-and-hold yielding -7.55% and the Base Strategy generating 18.55% returns alongside an exceptionally high

Sharpe ratio of 2.16. For gold and silver, predictive strategies enhanced returns and achieved better Sharpe ratios than the passive approach. Crude oil again acted as an exception: the buy-and-hold strategy produced an outstanding 35.48% return, whereas competing methods largely underperformed, with only the MAPE-filter strategy managing a positive outcome (10.89%). This discrepancy likely reflects that oil price movements were driven by nonlinear, unusual dynamics (geopolitical shocks, supply chain anomalies) to which the machine learning algorithms could not adequately respond. The cryptocurrency markets in 2022 experienced severe corrections. Buy-and-hold returns for Bitcoin (-85.01%), Ethereum (-120.93%), and Litecoin (-101.92%) incurred substantial losses. Predictive strategies, however, substantially mitigated losses, particularly for Litecoin, where the Base Strategy realized a 49.07% return with a Sharpe ratio of 2.17. Active strategy construction also generated profits for Bitcoin, although for Ethereum it could not fully offset the negative trend. Predictive strategies also produced more favorable outcomes in currency markets. The EUR/USD buy-and-hold return was -7.65%, while the Base Strategy (MAPE-based) achieved 7.91% with a Sharpe ratio of 2.57. For GBP/USD and AUD/USD, results were more mixed, but the passive approach still underperformed. Overall, in 2022, machine learning-based predictive strategies demonstrated their value, particularly in managing crisis-like, extreme market conditions. Across equities, cryptocurrencies, precious metals, and multiple currency pairs, they were able to reduce losses and, in many cases, deliver positive returns. Improvements in Sharpe ratios clearly indicate that these models manage risk more effectively. The oil market, however, remained a critical challenge, as investment approaches based on predictive models consistently underperformed in this asset class.

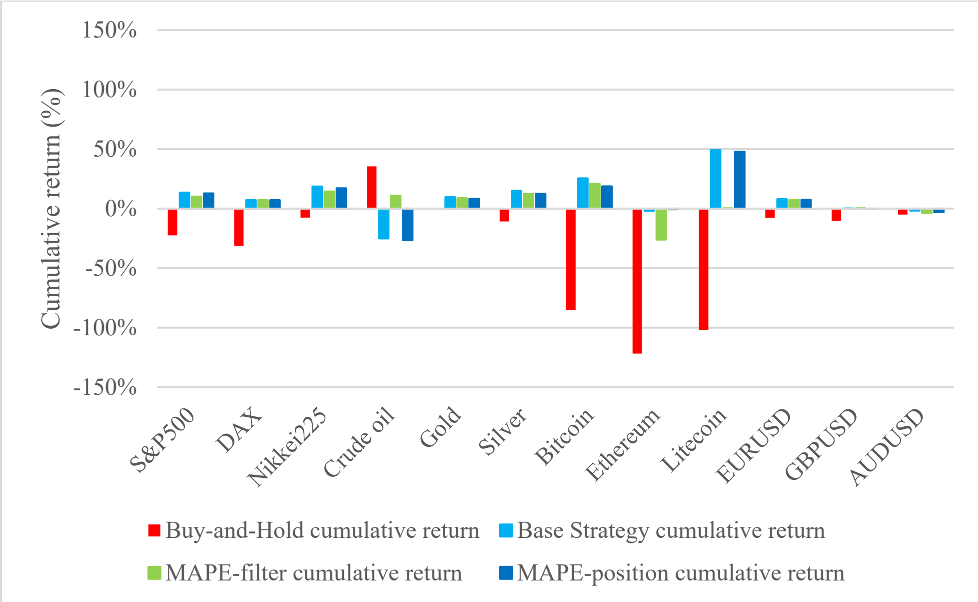

Figure 5: Comparison of the 4 trading strategies for the Russian–Ukrainian conflict period (2022)

Source: own editing

Across the three years under review, it is generally observed that machine learning-based active strategies systematically outperformed the buy-and-hold approach in terms of both returns and risk-adjusted performance. This trend was particularly pronounced in equity and cryptocurrency markets. The largest differences occurred in years with elevated market volatility (e.g., 2020 and 2022), supporting the conclusion that machine learning strategies are more effective at navigating turbulent environments. A notably important exception, however, is crude oil, which produced weak results for active strategies in every year. This may be attributed to the unique characteristics of commodity markets, the difficult-to-predict impacts of fundamental events (geopolitics, OPEC decisions), and the fact that oil price movements often occur suddenly within narrow time windows. Overall, the broad application of machine learning strategies can be considered successful, although certain asset classes (e.g., oil) may require further model development or combination with other predictive tools. Accordingly, predictive models represent a valuable supplement to investment decision-making, but they do not replace careful market interpretation. This highlights that while artificial intelligence holds significant potential, no single model provides a universal solution for all market conditions.

Beyond evaluating individual asset performance, I also experimented with portfolios of varying weightings, which are detailed in the following section.

Figure 6 presents the cumulative returns in 2018 for equally weighted portfolios across four different asset classes (each asset included with a 1/3 weight), broken down by strategy. The buy-and-hold approach underperforms in all portfolios, particularly in the cryptocurrency portfolio, which suffered a drastic -75.13% loss. In contrast, the Base Strategy significantly improves performance, especially for the cryptocurrency portfolio, achieving a cumulative return of 100.54%, thereby fully outperforming the passive investment approach. Models based on the MAPE-filter and MAPE-position strategies exhibit more stable, conservative results: for the commodity portfolio, both approaches produced around 12% returns, surpassing buy-and-hold. Improvements are also observed for the equity and currency portfolios when using active strategies, particularly those based on MAPE. In the cryptocurrency portfolio, the MAPE-filter strategy delivers more modest but still positive performance (15.21%), while the MAPE-position strategy achieves an outstanding 89.91% return, which is remarkable given the extremely volatile environment. Overall, active, prediction-based strategies effectively manage high-risk market conditions (cryptocurrencies), while moderate gains can also be achieved in more conservative asset classes.

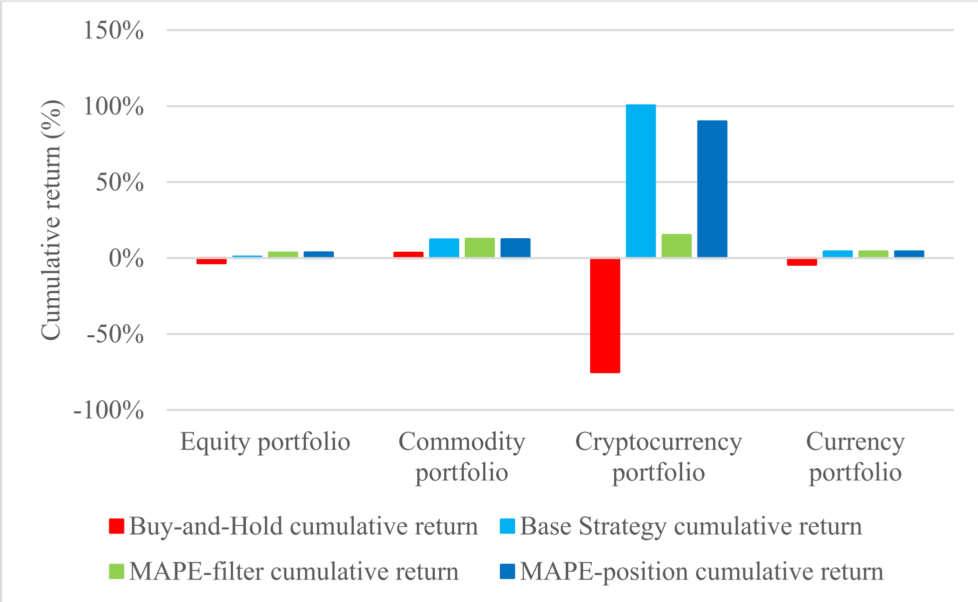

Figure 6: Cumulative semi-annual returns of equally weighted portfolios by asset class for the calm period (2018)

Source: own editing

Figure 7 presents the cumulative returns for various portfolios in 2020. Significant differences can be observed across asset classes and strategies. In the equity portfolio, the buy-and-hold strategy yielded a negative return (-7.51%), whereas the Base Strategy and the MAPE-position strategy showed substantial positive performance (37.74% and 36.90%), indicating that active, prediction-based approaches can outperform even in a crisis environment. For the commodity portfolio, although the buy-and-hold strategy achieved a positive return (11.17%), the Base and MAPE-position strategies resulted in negative returns (-2.11% and -1.92%), suggesting that these active models struggled to adapt effectively to market dynamics in this segment. The cryptocurrency portfolio delivered outstanding returns across all strategies, with the Base Strategy (120.96%) and MAPE-position strategy (107.37%) achieving exceptional performance, surpassing even the buy-and-hold approach (27.43%). In the currency markets, returns remained low under all approaches, yet prediction-based strategies still outperformed the passive buy-and-hold method. Overall, machine learning-based strategies were particularly effective in high-volatility markets (cryptocurrencies, equities), while their performance was moderate or occasionally underperforming in more stable or less predictable markets (commodities, currencies).

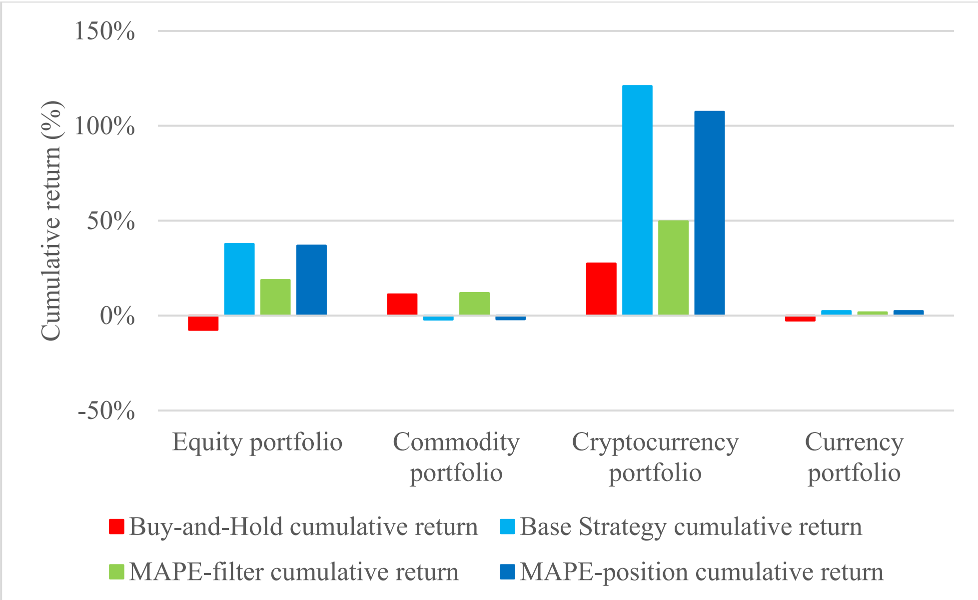

Figure 7: Cumulative semi-annual returns of equally weighted portfolios by asset class for the Covid19 period (2020)

Source: own editing

Based on the 2022 data (Figure 8), the buy-and-hold strategy performed poorly across all portfolios, particularly in the cryptocurrency portfolio, which experienced an extreme cumulative loss of -103.05%. In contrast, the Base Strategy consistently outperformed the passive approach, especially in the equity market (13.05%) and the crypto market (24.23%), achieving positive returns despite high volatility. The MAPE-filter strategy produced more stable and moderate results, delivering returns above 10% for equities and commodity assets, while closing slightly negative in the cryptocurrency portfolio (-1.67%). The MAPE-position strategy also performed strongly in both the equity (12.29%) and cryptocurrency portfolios (21.99%), confirming that position management based on prediction errors can effectively mitigate risk in turbulent market conditions. For the commodity portfolio, both the Base and MAPE-position strategies showed slight losses, indicating that these models were less capable of accurately tracking the market movements in this segment. In the currency markets, all approaches yielded similar, modest positive returns (approximately 1.4–2.1%), reflecting low volatility and limited predictability. Overall, active, prediction-based strategies in 2022 again outperformed the buy-and-hold approach, particularly in equity and cryptocurrency markets, even under extreme market stress conditions.

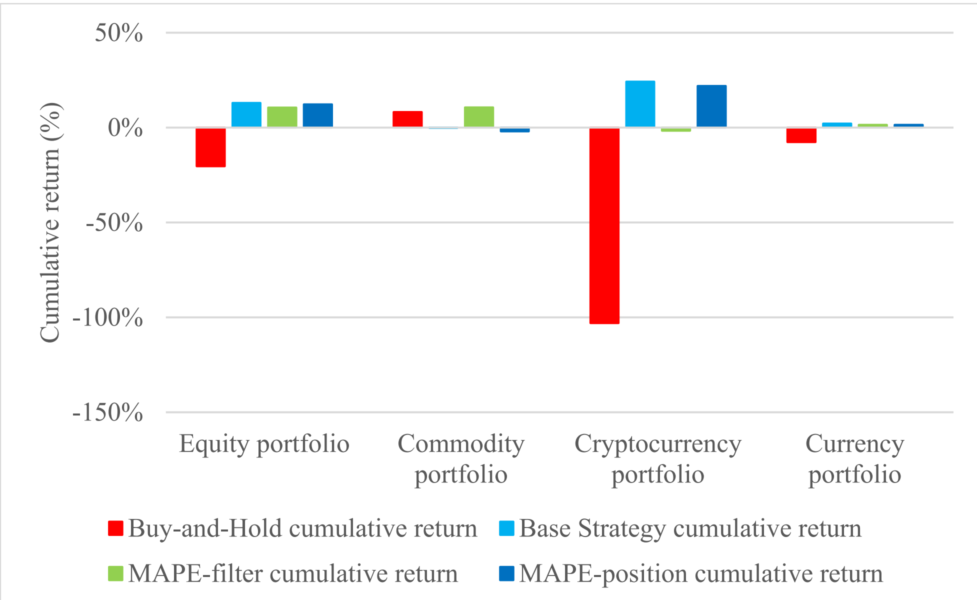

Figure 8: Cumulative semi-annual returns of equally weighted portfolios by asset class for the Russian-Ukrainian conflict period (2022)

Source: own editing

Figure 9 presents the cumulative returns of equally weighted portfolios (each asset assigned a 1/12 weight) across three different years, based on four distinct strategic approaches. The buy-and-hold strategy recorded significant losses in 2018 and 2022 (-19.86% and -30.70%, respectively), while achieving a moderate 7.12% return in 2020. In contrast, the Base Strategy generated positive cumulative returns in all examined years, notably outperforming in 2020 (39.75%) and strongly exceeding the buy-and-hold performance in 2018 (29.52%). The MAPE-filter-based strategy delivered more conservative but generally more stable results, closing 2018 with an 8.98% return, 2020 with 20.52%, and 2022 with 5.26% positive returns. The MAPE-position approach also significantly outperformed the passive strategy, particularly in 2020 and 2018, achieving cumulative returns of +36.17% and +27.60%, respectively. These findings indicate that filtering and position management based on prediction errors (MAPE) can help mitigate negative market impacts and enhance performance stability under varying environmental conditions.

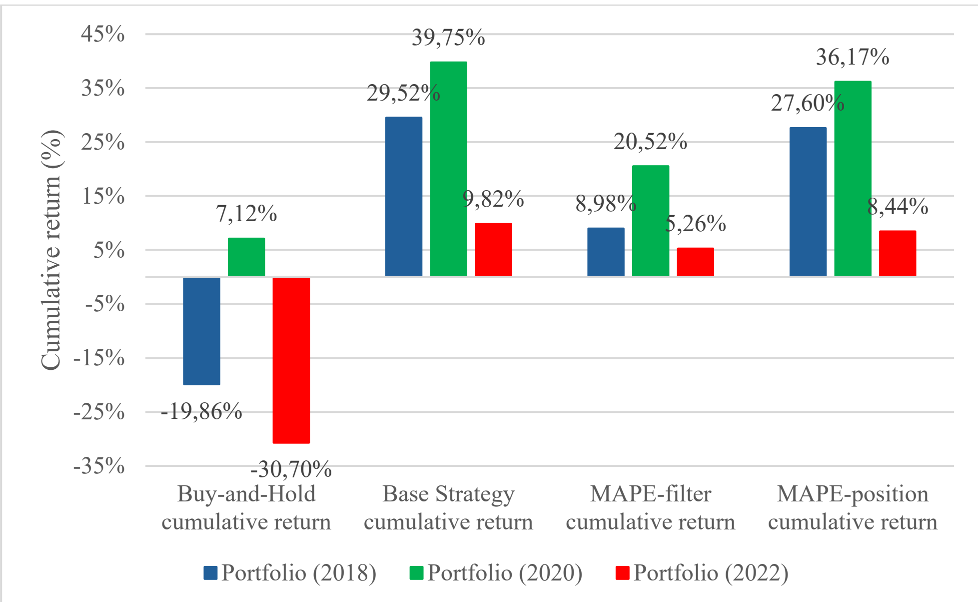

Figure 9: Semi-annual cumulative returns of equally weighted portfolios composed of 12 assets for the calm period, the Covid19 period, and the Russia-Ukraine conflict period (2018, 2020, and 2022)

Source: own editing

5. CONCLUSIONS AND RECOMMENDATIONS

The objective of the research was to explore the applicability of machine learning algorithms, particularly deep neural networks, for predicting the prices of financial instruments in different market environments and product types. The study focused on the volatility sensitivity of predictive models, the reliability of predictions, and their practical applicability. The thesis examined the topic based on two main objectives. The first focuses on the relationship between the performance of predictive models and market volatility, as well as the applicability of individual models in different economic cycles. The second objective focuses on the practical performance of trading strategies supported by machine learning models. The analysis paid particular attention to the modeling of stock indices, commodity market products, cryptocurrencies, and currency pairs, as their volatility, liquidity, and regulation vary greatly. Based on the empirical analyses conducted during the research, well-founded conclusions can be drawn from several perspectives, which I will detail below.

C1: The primary objective of my research is to determine the extent to which different neural deep learning models can be generalized, i.e., whether they are capable of achieving outstanding predictive performance in different crisis situations.

K1: What relationship can be demonstrated between the volatility of financial instruments and the predictive performance of price forecasting models?

Based on regression models, it can be clearly demonstrated that volatility, in this case the trend and seasonality-free (CV_STL) relative standard deviation, has a significant positive correlation with the forecast error (MAPE). According to my results, the accuracy of forecasts deteriorates in all model categories examined as market volatility increases, which is particularly noticeable in the case of cryptocurrencies, but can also be observed in commodity market products. This is consistent with the findings of several studies that emphasize the difficulties of forecasting nonlinear and non-stationary data (Ouyang et al., 2021; Zhang et al., 2023; Avinash et al., 2024). Deep learning can partially address these challenges, but sudden shifts in turbulent markets remain a problem in terms of prediction reliability.

K2: What kind of forecasting distortions are caused by the extreme price movements observed during periods of crisis, and how do different types of algorithms respond to this?

The crisis periods of 2020 and 2022 clearly demonstrated that models respond differently to extreme price movements. GRU models, for example, proved to be surprisingly stable, while the performance of hybrid architectures containing RNNs (e.g., RNN-LSTM) deteriorated significantly, meaning that the crisis periods not only caused an increase in prediction errors but also exacerbated performance differences between models. This confirms the findings reported in the literature that overly complex models are more susceptible to instability (Livieris et al., 2020) and also highlights that forecasting algorithms are sensitive to extreme price movements caused by market shocks regardless of their level of sophistication, so it is of paramount importance to take these factors into account when developing forecasting methods (Mari and Mari, 2023 and 2025; Sivakumar, 2025).

K3: What role does hybrid model architecture play in predictive performance change in a volatile market environment?

Based on the results of this paper, it can be concluded that the architecture of predictive models, especially hybrid deep learning structures, plays a significant role in the development of forecasting performance, particularly in volatile market environments. Empirical studies have shown that models using

RNN components as the first layer often perform worse than other architectures, especially in the case of cryptocurrencies. This result is partly consistent with the findings of Liang et al. (2022) and Ouyang et al. (2021), who pointed out that more complex, attention-based or Transformer-type architectures can provide more robust performance in extreme market environments. At the same time, the simpler LSTM- or GRU-based models examined in the dissertation proved to be more stable than RNN modules in several cases, especially in hybrid structures, which supports the importance of fine-tuning hybrid architectures. The internal structure of the models is therefore not just a matter of technical preference, but a key element in managing volatility.

C2: The second objective of my thesis is to examine the extent to which machine learning models can improve the trading performance of financial instruments compared to the traditional buy-and-hold strategy.

K4: How well can trading strategies based on machine learning predictions exploit market anomalies, as opposed to the efficient market theory?

Backtesting of rule-based, prediction-driven trading strategies confirmed that they outperformed the buy-and-hold benchmark return in several periods and for several product types. Versions equipped with a MAPE filter enabled additional risk management, as they were able to filter out positions with low prediction certainty. My results confirm the findings of Viéitez et al. (2024) and Ju et al. (2024), who argue that, in addition to the predictive capabilities of the models, their adaptive use, for example in the form of position filtering, greatly improves practical performance. In the case of the Sharpe ratio, it was observed that the predictive strategy generally showed a higher value, indicating a better risk-return ratio in terms of return volatility. This finding contradicts the efficient market hypothesis (EMH), which states that it is not possible to systematically outperform the market in the long run by taking all available information into account (Fama, 1970).

K5: What effect does the volatility of different asset classes (e.g., stocks, cryptocurrencies, commodities, currency pairs) have on the performance of machine learning-supported trading strategies?

The results show that the unpredictability of cryptocurrencies and the stability of currency pairs differ significantly, which has a direct impact on trading results. Product-specific analyses have therefore highlighted that volatility affects strategy performance in different ways. In the case of cryptocurrencies, the MAPE filter was particularly effective in reducing losses, while in foreign exchange markets, the stability of the Base Strategy was outstanding. In stock markets, the relative performance of the strategies proved to be more period-dependent, which is consistent with the findings in the literature that asset-specific volatility requires a different modeling approach (Yu et al., 2023; Aydogan-Kilic and Selcuk-Kestel, 2023). The results support the need for product-specific parameterization to achieve maximum efficiency. Another important conclusion is that, in the case of equally weighted portfolios, the performance of prediction-driven strategies was not consistent across all market segments. For commodity and currency market instruments, the differences between the strategic variants were smaller, while for cryptocurrencies, there were drastic differences. This reaffirms that the effectiveness of a strategy and model can only be interpreted in relation to a given product type and period, so model selection and parameter tuning play a decisive role in practical effectiveness.

K6: What differences can be observed in the performance of machine learning-based strategies during different economic cycles?

The performance of trading strategies varied significantly across different market environments. In the calm market environment of 2018, almost all model-based strategies generated stable profits. In contrast, in 2020 and 2022, predictive performance declined due to increased volatility. However, MAPE-filtered systems reduced the risk of losses, confirming the practical relevance of Ouyang et al. (2021) attention-based volatility management strategies and highlighting the importance of adaptive decision-making mechanisms. These findings reinforce the view in the literature (Kang et al., 2025) that prediction systems must adapt to changes in economic regimes. The performance of strategies thus depends not only on the accuracy of the model, but also on its volatility-sensitive application. One of the most important conclusions of the research is therefore that machine learning-based trading systems can only deliver stable performance in crisis situations if they also have built-in prediction validation mechanisms and adaptive filtering.

The empirical results of the dissertation confirm that machine learning models are effective prediction tools, but their performance depends significantly on volatility, market regime, and asset type. The excess returns and risk reduction achieved through the use of prediction-based strategies demonstrate that such systems can be valuable additions to investment decision support, as they offer not only theoretical but also practical advantages, provided that they adequately integrate the measurement of forecast uncertainty and adaptation to market volatility.

6. NEW SCIENTIFIC RESULTS

I. **I have verified through joint analysis of several different product types (stocks, commodity instruments, currency pairs, cryptocurrencies) that volatility has a significant impact on the predictive performance (MAPE) of neural network models, regardless of the methodology used.**

During my research, I examined different asset classes (stock indices, commodity products, currency pairs, and cryptocurrencies) in three different periods. Based on the results, there was a significant positive correlation between trend- and seasonality-adjusted relative volatility and the mean absolute percentage error (MAPE) in all cases. The panel regression models showed high R^2 values, which supports the fact that volatility is not only statistically significant but also has a strong explanatory power in terms of changes in predictive performance. This effect was independent of the asset class type and the chosen model architecture, meaning that volatility acts as a distorting factor in all cases.

II. **I have demonstrated using empirical tools that the level of market volatility influences the performance difference between models, which is noticeable during periods of high volatility.**

During the 2018 ("calm") period, the forecast errors of the individual models were pretty much the same. Most models (RNN, LSTM, GRU, and their hybrids) produced similar MAPE values, so model selection wasn't that important. During periods of lower volatility, the performance of the models converged, reducing the significance of model selection. In contrast, during the 2020 crisis period, the differences between the models increased significantly: the standard deviation of prediction errors more than doubled. The studies also revealed that higher volatility levels increase not only MAPE but also performance variance, which is a particularly important factor in model application.

III. **During the periods examined, the RNN architecture used in the first hidden layer of hybrid models has a performance-degrading effect, regardless of which other algorithm it is combined with.**

Based on the regression results for the 2018 period, it is clear that hybrid models based on RNN (RNN-GRU or RNN-LSTM) are characterized by higher volatility sensitivity and lower predictive performance than hybrid models that do not contain an RNN component in the first layer. This suggests that RNN structures have more difficulty handling densely noisy, trend-dependent time series, especially in cases of high volatility. The results also show that not all model combinations lead to performance gains and that the internal structure of the network, especially the structure of the first layers, plays a critical role in prediction accuracy.

IV. **I have demonstrated that, during the period under review, GRU-based forecasting models are less sensitive to increases in relative volatility than RNN- or LSTM-based architectures, which offers a novel approach to examining the relationship between model specification and market uncertainty.**

Based on the panel regression analyses presented in the dissertation, relative volatility was significantly positively correlated with the mean absolute percentage error (MAPE), regardless of the period and asset class. At the same time, I also found that the model response to volatility is not uniform: architectures were distorted in different ways and to different degrees by increased market uncertainty. GRU-based models showed lower sensitivity, while the performance of RNN-based hybrid architectures deteriorated significantly in the case of high volatility. This result combines the quantitative impact of the volatility environment with the quality of the model specification in a novel way and confirms that prediction risk depends not only on the current state of the market but also on the chosen model structure. This enables a preliminary analysis of the volatility sensitivity profile of models, which can be an important practical guideline, for example, in the application of predictive trading systems. This finding is consistent with, but also further develops, the conclusions of the current literature, which discusses the interaction between volatility and models separately, but does not systematically compare the volatility sensitivity of different architectures.

V. **I have demonstrated that trading rules (position opening timing and position sizing) developed with dynamic consideration of the mean absolute percentage error (MAPE) significantly improve the risk-return ratio, especially in the case of highly volatile asset classes. This result provides empirical evidence against the validity of the weak form of the efficient market hypothesis.**

During the research, trading strategies were not only based on the direction of price movements predicted by the model, but also integrated the estimated prediction error (MAPE). This allowed the strategy to manage uncertainty on two levels. On the one hand, it excluded opening positions during periods when the reliability of predictions was low. On the other hand, it adapted directly to the performance of the models at the level of capital exposure by dynamically scaling position sizes. The results were particularly striking for cryptocurrencies, as the maximum drawdown for Litecoin and Ethereum decreased significantly, while the cumulative return and Sharpe ratio increased significantly compared to the traditional buy-and-hold strategy. This approach offers a new perspective on the use of predictive models, as it not only uses deterministic signals for trading decisions, but also explicitly incorporates model uncertainty into decision-making. This contributes to the development of an adaptive, self-reflective trading architecture that is sensitive to changes in market regimes. In practice, performance suggests that by taking prediction errors into account, it is possible to achieve statistically significant excess returns in certain market segments especially in asset classes characterized by high volatility and information asymmetry even when the weak form efficient market hypothesis holds.

7. SUMMARY

The emergence of complex and large data sets has caused significant technological and conceptual changes in the development and application of forecasting models over the past 20 years. Processing the vast amount of data would no longer be efficient using traditional methods, and machine learning has become an essential part of most predictive modelling industries. This is particularly true in the financial sector, where profitable operations seek to make the best use of the innovative tools available, which are much needed to explore non-linear relationships and patterns.

The dissertation examines two highly topical and practically significant financial prediction modeling problems. The first objective was to map the impact of volatility on prediction performance, while the second evaluated the effectiveness of trading strategies utilizing machine learning-based predictions in different market environments. The research relied on quantitative regression analyses, comparisons of deep learning architectures, and back-tested trading strategies, thus contributing new scientific knowledge to financial machine learning research in both theoretical and applied aspects. In the introduction, I presented the economic background and significance of the field, and also discussed the development path of predictive algorithms. Furthermore, I provided insight into the current state of this field of science and the latest trends. The topic I examined is quite popular among international researchers, and there are many publications to choose from to get a more detailed picture. I divided the systematic literature review into four main sections, in which I dealt in detail with the stock, cryptocurrency, commodity, and foreign exchange markets. The analysis of scientific publications provided guidance in understanding the characteristics of the models on the one hand, and in identifying various product-specific factors on the other. In the empirical part of my research, I used data from the most traded stock indices (S&P500, DAX, Nikkei225), commodity market products (crude oil, gold, silver), cryptocurrencies (Bitcoin, Ethereum, Litecoin) and currency pairs (EUR/USD, GBP/USD, AUD/USD) for the period from January 1, 2016 to June 30, 2022. This period was chosen in part because it includes the calm period (2018), the Covid19 (2020) and the war crisis (2022), and partly because cryptocurrencies are relatively new products compared to the others, so their price data covers a shorter period. Therefore, this seemed to be the optimal decision in terms of comparability. I collected the data from the website www.finance.yahoo.com, with the exception of cryptocurrencies, for which the data comes from www.coinmarketcap.com. The modeling was based on three deep learning algorithms (RNN, LSTM, GRU) and three hybrid methodologies developed from them (LSTM-GRU, RNN-LSTM, RNN-GRU) developed from them. I presented the differences between actual and estimated prices using the MAPE indicator. I divided the results chapter into three main sections. In the first section, I compared the models examined by product type, model, and period. The performance analysis of the methods was essential for further investigations, which focused on volatility and real trading strategies. The first objective of the thesis was to explore the relationship between the different volatility structures of financial instruments and the accuracy of time series-based machine learning models. Experiments conducted on different product classes (stocks, commodity market instruments, cryptocurrencies, and currency pairs) confirmed that there is a strong positive correlation between relative volatility and prediction error, which was significantly supported by regression models run on different architectures. The relationship was present in all periods examined, but proved to be particularly strong in the crisis periods of 2020 and 2022. Another important finding of the research is that the performance of the models is highly dependent on the period and the asset. During the calm period (2018), the differences were minimized, while during the crisis periods (2020, 2022), model selection became a critical factor. Hybrid architectures (RNN-GRU, LSTM-GRU) performed better in most cases, but not all combinations were beneficial, especially the use of RNN as the first layer, which degraded performance. The research thus directly confirmed the assumption that neither volatility nor model architecture can be treated as independent factors when examining predictive performance. The second objective of the research was to examine whether trading strategies built on machine learning models are capable of systematically outperforming the passive investment approach (buy-and-hold), especially in environments with varying volatility. The paper examined backtested strategies in equally weighted portfolios over three years (2018, 2020, 2022) and across four different asset classes. Methods based on the predictive performance of machine learning models (Base Strategy, MAPE-filter, and MAPE-position) consistently improved the risk-return ratio, which was most evident in lower Sharpe ratios and drawdowns. In the case of cryptocurrencies, machine learning-based strategies were most prominent in 2022, while losses were significantly reduced. In line with the trading strategy literature discussed in Section 2.8, the results of this paper show that machine learning not only offers predictive accuracy but also has decision support potential. The difference is that the present study explicitly quantified the volatility sensitivity of the strategies, which provides a new perspective on the evaluation of the applicability of machine learning in an investment environment. The success of the strategies calls into question the practical validity of the efficient market hypothesis (EMH), especially if the models used are able to adapt to market regimes, as this research has demonstrated on several occasions.

Overall, the results of this paper contribute to the understanding of the volatility sensitivity of machine learning models and support the notion that predictive performance depends not only on the structure of the model but also a large extent on the characteristics of the market environment. Based on the literature review, it can be stated that although numerous publications deal with volatility forecasting, few studies specifically analyze the impact of volatility on predictive performance, especially when examining multiple asset classes simultaneously. The examination of trading strategies based on predictive models further reinforces the practical economic value of the thesis. The research draws attention to the diversity and excellent results of machine learning in the financial sector and emphasizes that the application of the latest technologies is essential for continuous development and profitable operation.

8. REFERENCES

1. Atsalakis, G. S., & Valavanis, K. P. (2009). Surveying stock market forecasting techniques–Part II: Soft computing methods. Expert Systems with applications, 36(3), 5932-5941. doi: https://doi.org/10.1016/j.eswa.2008.07.006
2. Avinash, G., Ramasubramanian, V., Ray, M., Paul, R. K., Godara, S., Nayak, G. H. H., ... & Iquebal, M. A. (2024). Price Forecasting of TOP (Tomato, Onion and Potato) Commodities using Hidden Markov-based Deep Learning Approach. Statistics and Applications, 22(2), 1-28.
3. Aydogan-Kilic, D., & Selcuk-Kestel, A. S. (2023). Modification of hybrid RNN-HMM model in asset pricing: univariate and multivariate cases. Applied Intelligence, 53(20), 23812-23833. doi: https://doi.org/10.1007/s10489-023-04762-7
4. Bai, Y., Xie, J., Liu, C., Tao, Y., Zeng, B., & Li, C. (2021). Regression modeling for enterprise electricity consumption: A comparison of recurrent neural network and its variants. International Journal of Electrical Power & Energy Systems, 126, 106612. doi: https://doi.org/10.1016/j.ijepes.2020.106612
5. Borovkova, S., & Tsiamas, I. (2019). An ensemble of LSTM neural networks for high-frequency stock market classification. Journal of Forecasting. doi: https://doi.org/10.1002/for.2585
6. Brockwell, P. J., & Davis, R. A. (Eds.). (2002). Introduction to time series and forecasting. New York, NY: Springer New York.
7. Büyükşahin, Ü. Ç., & Ertekin, Ş. (2019). Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition. Neurocomputing, 361, 151-163. doi: https://doi.org/10.1016/j.neucom.2019.05.099
8. Chen, M. Y., & Chen, B. T. (2015). A hybrid fuzzy time series model based on granular computing for stock price forecasting. Information Sciences, 294, 227-241. doi: https://doi.org/10.1016/j.ins.2014.09.038
9. Chen, Y., & Hao, Y. (2017). A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction. Expert Systems with Applications, 80, 340-355. doi: https://doi.org/10.1016/j.eswa.2017.02.044

10. Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition procedure based on loess. Journal of Official Statistics, 6(1), 3–73.

11. Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor Psychology and Security Market Under- and Overreactions. The Journal of Finance, 53(6), 1839–1885. doi: https://doi.org/10.1111/0022-1082.00077

12. De BONDT, W. F. M., & THALER, R. (1985). Does the Stock Market Overreact? The Journal of Finance, 40(3), 793–805. doi: 10.1111/j.1540-6261.1985.tb05004.x

13. Duarte, J. B. D., Sarmiento, L. H. T., & Juárez, K. J. S. (2017). Evaluation of the effect of investor psychology on an artificial stock market through its degree of efficiency. Contaduría y Administración, 62(4), 1361-1376. doi: https://doi.org/10.1016/j.cya.2017.06.014

14. Efendi, R., Arbaiy, N., & Deris, M. M. (2018). A new procedure in stock market forecasting based on fuzzy random auto-regression time series model. Information Sciences, 441, 113-132. doi: https://doi.org/10.1016/j.ins.2018.02.016

15. Fama, E. F. (1965). Random Walks in Stock Market Prices. Financial Analysts Journal, 21(5), 55–59.

16. Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance, 25(2), 383. doi: https://doi.org/10.2307/2325486

17. Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. Journal of financial economics, 49(3), 283-306. doi: https://doi.org/10.1016/S0304-405X(98)00026-9

18. Faru, S. H., Waititu, A., & Nderu, L. (2023). A Hybrid Neural Network Model Based on Transfer Learning for Forecasting Forex Market. Journal of Data Analysis and Information Processing, 11(2), 103-120. doi: 10.4236/jdaip.2023.112007

19. He, H., Gao, S., Jin, T., Sato, S., & Zhang, X. (2021). A seasonal-trend decomposition-based dendritic neuron model for financial time series prediction. Applied Soft Computing, 108, 107488. doi: https://doi.org/10.1016/j.asoc.2021.107488

20. Hoseinzade, E., & Haratizadeh, S. (2019). CNNpred: CNN-based stock market prediction using a diverse set of variables. Expert

Systems with Applications, 129, 273-285. doi: https://doi.org/10.1016/j.eswa.2019.03.029

21. Islam, M. S., & Hossain, E. (2021). Foreign exchange currency rate prediction using a GRU-LSTM hybrid network. Soft Computing Letters, 3, 100009. doi: https://doi.org/10.1016/j.socl.2020.100009
22. Ju, C., Shen, Q., & Ni, X. (2024). Leveraging LSTM Neural Networks for Stock Price Prediction and Trading Strategy Optimization in Financial Markets. Applied and Computational Engineering, 112, 47-53. doi: https://doi.org/10.54254/2755-2721/2024.17906
23. Kang, M., Hong, J., & Kim, S. (2025). Harnessing technical indicators with deep learning based price forecasting for cryptocurrency trading. Physica A: Statistical Mechanics and its Applications, 130359. doi: https://doi.org/10.1016/j.physa.2025.130359
24. Kim, H. Y., & Won, C. H. (2018). Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. Expert Systems with Applications, 103, 25-37. doi: https://doi.org/10.1016/j.eswa.2018.03.002
25. Kumbure, M. M., Lohrmann, C., Luukka, P., & Porras, J. (2022). Machine learning techniques and data for stock market forecasting: A literature review. Expert Systems with Applications, 197, 116659. doi: https://doi.org/10.1016/j.eswa.2022.116659
26. Liang, Y., Lin, Y., & Lu, Q. (2022). Forecasting gold price using a novel hybrid model with ICEEMDAN and LSTM-CNN-CBAM. Expert Systems with Applications, 206, 117847. doi: https://doi.org/10.1016/j.eswa.2022.117847
27. Livieris, I. E., Pintelas, E., & Pintelas, P. (2020). A CNN–LSTM model for gold price time-series forecasting. Neural Computing and Applications. doi:10.1007/s00521-020-04867-x
28. Lo, A. W. (2004). The Adaptive Markets Hypothesis. The Journal of Portfolio Management, 30(5), 15–29. doi: 10.3905/jpm.2004.442611
29. Lohrmann, C., & Luukka, P. (2019). Classification of intraday S&P500 returns with a Random Forest. International Journal of Forecasting, 35(1), 390-407.Lohrmann, C., & Luukka, P. (2018). Classification of intraday S&P500 returns with a Random Forest. International Journal of Forecasting. doi: 10.1016/j.ijforecast.2018.08.004

30. Malkiel, B., Mullainathan, S., & Stangle, B. (2005). Market Efficiency versus Behavioral Finance. Journal of Applied Corporate Finance, 17(3), 124–136. doi: 10.1111/j.1745-6622.2005.00053.x

31. Maqsood, H., Mehmood, I., Maqsood, M., Yasir, M., Afzal, S., Aadil, F., ... & Muhammad, K. (2020). A local and global event sentiment based efficient stock exchange forecasting using deep learning. International Journal of Information Management, 50, 432-451. doi: https://doi.org/10.1016/j.ijinfomgt.2019.07.011

32. Mari, C., & Mari, E. (2023). Deep learning based regime-switching models of energy commodity prices. Energy Systems, 14(4), 913-934. doi: https://doi.org/10.1007/s12667-022-00515-6

33. Mari, C., & Mari, E. (2025). Stochastic DNN-based models meet hidden Markov models: A challenge on natural gas prices at the Henry Hub. Neural Computing and Applications, 1-20. doi: https://doi.org/10.1007/s00521-025-11286-3

34. Nabipour, M., Nayyeri, P., Jabani, H., Shahab, S., & Mosavi, A. (2020). Predicting stock market trends using machine learning and deep learning algorithms via continuous and binary data; a comparative analysis on the Tehran stock exchange. IEEE Access, 1–1. doi: 10.1109/access.2020.3015966

35. Ortu, M., Uras, N., Conversano, C., Bartolucci, S., & Destefanis, G. (2022). On technical trading and social media indicators for cryptocurrency price classification through deep learning. Expert Systems with Applications, 198, 116804. doi: https://doi.org/10.1016/j.eswa.2022.116804

36. Ouyang, Z. S., Yang, X. T., & Lai, Y. (2021). Systemic financial risk early warning of financial market in China using Attention-LSTM model. The North American Journal of Economics and Finance, 56, 101383. doi: https://doi.org/10.1016/j.najef.2021.101383

37. Patel, P. B., & Marwala, T. (2006, October). Forecasting closing price indices using neural networks. In 2006 IEEE international conference on systems, man and cybernetics (Vol. 3, pp. 2351-2356). IEEE. doi: 10.1109/ICSMC.2006.385214

38. Reston Filho, J. C., Affonso, C. D. M., & de Oliveira, R. C. (2014). Energy price prediction multi-step ahead using hybrid model in the Brazilian market. Electric power systems research, 117, 115-122. doi: https://doi.org/10.1016/j.epsr.2014.08.006

39. Sedighi, M., Jahangirnia, H., Gharakhani, M., & Farahani Fard, S. (2019). A novel hybrid model for stock price forecasting based on metaheuristics and support vector machine. Data, 4(2), 75. doi: https://doi.org/10.3390/data4020075

40. Shiller, R. J. (2003). From efficient markets theory to behavioral finance. Journal of economic perspectives, 17(1), 83-104. doi: https://doi.org/10.1257/089533003321164967

41. Sivakumar, G. (2025). HMM-LSTM Fusion Model for Economic Forecasting. arXiv preprint arXiv:2501.02002. doi: https://doi.org/10.48550/arXiv.2501.02002

42. Song, Y., Lee, J. W., & Lee, J. (2019). A study on novel filtering and relationship between input-features and target-vectors in a deep learning model for stock price prediction. Applied Intelligence, 49, 897-911. doi: https://doi.org/10.1007/s10489-018-1308-x

43. Timmermann, A., & Granger, C. W. (2004). Efficient market hypothesis and forecasting. International Journal of forecasting, 20(1), 15-27. doi: https://doi.org/10.1016/S0169-2070(03)00012-8

44. Vidal, A., & Kristjanpoller, W. (2020). Gold volatility prediction using a CNN-LSTM approach. Expert Systems with Applications, 157, 113481. doi: https://doi.org/10.1016/j.eswa.2020.113481

45. Viéitez, A., Santos, M., & Naranjo, R. (2024). Machine learning Ethereum cryptocurrency prediction and knowledge-based investment strategies. Knowledge-Based Systems, 299, 112088. doi: https://doi.org/10.1016/j.knosys.2024.112088

46. Wei, L. Y., Chen, T. L., & Ho, T. H. (2011). A hybrid model based on adaptive-network-based fuzzy inference system to forecast Taiwan stock market. Expert Systems with Applications, 38(11), 13625-13631. doi: https://doi.org/10.1016/j.eswa.2011.04.127

47. Weng, B., Ahmed, M. A., & Megahed, F. M. (2017). Stock market one-day ahead movement prediction using disparate data sources. Expert Systems with Applications, 79, 153-163. doi: https://doi.org/10.1016/j.eswa.2017.02.041

48. Xiao, H., Chen, Z., Cao, R., Cao, Y., Zhao, L., & Zhao, Y. (2022). Prediction of shield machine posture using the GRU algorithm with adaptive boosting: A case study of Chengdu Subway project. Transportation Geotechnics, 37, 100837. doi: https://doi.org/10.1016/j.trgeo.2022.100837

49. Yeh, C. Y., Huang, C. W., & Lee, S. J. (2011). A multiple-kernel support vector regression approach for stock market price forecasting. Expert Systems with Applications, 38(3), 2177-2186. doi: https://doi.org/10.1016/j.eswa.2010.08.004

50. Yu, Y., Lin, Y., Hou, X., & Zhang, X. (2023). Novel optimization approach for realized volatility forecast of stock price index based on deep reinforcement learning model. Expert Systems with Applications, 120880. doi: https://doi.org/10.1016/j.eswa.2023.120880

51. Zhang, J., Ye, L., & Lai, Y. (2023). Stock price prediction using CNN-BiLSTM-Attention model. Mathematics, 11(9), 1985. doi: https://doi.org/10.3390/math11091985

52. Zhao, L., Li, Z., Qu, L., Zhang, J., & Teng, B. (2023). A hybrid VMD-LSTM/GRU model to predict non-stationary and irregular waves on the east coast of China. Ocean Engineering, 276, 114136. doi: https://doi.org/10.1016/j.oceaneng.2023.114136

53. Zolfaghari, M., & Gholami, S. (2021). A hybrid approach of adaptive wavelet transform, long short-term memory and ARIMA-GARCH family models for the stock index prediction. Expert Systems with Applications, 182, 115149. doi: https://doi.org/10.1016/j.eswa.2021.115149

9. PUBLICATIONS LIST (MTMT)

Bareith, T., Tatay, T., & Vancsura, L. (2023). Árupiaci termékek árfolyamelőrejelzése korszerű eszközökkel átmeneti gazdasági időszakban. Előadás. XVII. Soproni Pénzügyi Napok: „A mesterséges intelligencia szerepe a fenntartható gazdasági döntésekben" pénzügyi, adózási és számviteli szakmai és tudományos konferencia. Sopron: Soproni Egyetem Lámfalussy Sándor Közgazdaságtudományi Kar, 2023. október 11-13.

Vancsura, L., Tatay, T., & Bareith, T. (2023). Evaluating the effectiveness of modern forecasting models in predicting commodity futures prices in volatile economic times. Risks, 11(2), 27. doi: https://doi.org/10.3390/risks11020027

Vancsura, L., & Bareith, T. (2023). Analysis of the performance of predictive models during Covid-19 and the Russian-Ukrainian war. Public Finance Quarterly= Pénzügyi Szemle, 69(2), 118-132. doi: https://doi.org/10.35551/PFQ_2023_2_7

Bareith, T., Tatay, T., & Vancsura, L. (2024). Navigating Inflation Challenges: AI-Based Portfolio Management Insights. Risks, 12(3), 46. doi: https://doi.org/10.3390/risks12030046

Vancsura, L., Tatay, T., & Bareith, T. (2024). Investigating the Role of Activation Functions in Predicting the Price of Cryptocurrencies during Critical Economic Periods. Virtual Economics, 7(4), 64–91. https://doi.org/10.34021/ve.2024.07.04(4)

Vancsura, L., Tatay, T., & Bareith, T. (2025). Navigating AI-Driven Financial Forecasting: A Systematic Review of Current Status and Critical Research Gaps. Forecasting, 7(3), 36. https://doi.org/10.3390/forecast7030036