Research in International Business and Finance 66 (2023) 102052

Contents lists available at ScienceDirect

Research in International Business and Finance

journal homepage: www.elsevier.com/locate/ribaf

# Application of machine learning in algorithmic investment strategies on global stock markets☆

Jan Grudniewicz [a], Robert Ślepaczuk [b,*]

[a] *Quantitative Finance Research Group, Faculty of Economic Sciences, University of Warsaw, Ul. Długa 44/50, 00-241 Warsaw, Poland*
[b] *Quantitative Finance Research Group, Department of Quantitative Finance, Faculty of Economic Sciences, University of Warsaw, Ul. Długa 44/50, 00-241 Warsaw, Poland*

ARTICLE INFO

*JEL Classification:*
C4
C14
C45
C53
C58
G11
G17

*Keywords:*
Algorithmic investment strategies
Machine learning
Neural networks
Regression trees
Random forests
Support vector machine
Technical analysis
Equity stock indices
Developed and emerging markets
Information ratio

ABSTRACT

The research undertakes the subject of machine learning based algorithmic investment strategies. Several technical analysis indicators were employed as inputs to machine learning models such as Neural Networks, K Nearest Neighbor, Regression Trees, Random Forests, Naïve Bayes classifiers, Bayesian Generalized Linear Models, and Support Vector Machines. Models were used to generate trading signals on WIG20, DAX, S&P500, and selected CEE indices in the period between 2002-01–01 and 2023–03–31. Strategies were compared with each other and with the benchmark buy-and-hold strategy in terms of achieved levels of risk and return. Sensitivity analysis was used to assess the quality of the estimation on independent subsets. The findings of the study showed that algorithmic strategies outperformed passive strategies in terms of risk-adjusted returns and that for the analyzed indices, Linear Support Vector Machine and Bayesian Generalized Linear Model were the best-performing models. The Linear Support Vector Machine was chosen as the model that, on average, produced the best results using a more thorough rank approach based on the outcomes for all examined models and indices.

## 1. Introduction

Investing is the process of allocating capital to obtain future financial benefits. The investor expects the future cash flows from an investment to exceed its initial value. Investments in financial instruments are one of the methods of investing used by both individuals and institutional entities. One type of these instruments is a security traded on stock exchanges. The main stock exchange index is an indicator measuring the value of shares of the largest companies traded on a given stock exchange. The construction of investment portfolios reflecting the behavior of stock exchange indices is a method of passive capital management.

An alternative to passive investing is active capital management. Investors can actively select equity instruments in their

☆ We want to thank for the inspiring comments of the participants of the 19th Quantitative Finance Research Group and Data Science Lab Research Seminar at the Faculty of Economic Sciences at the University of Warsaw.

* Corresponding author.
*E-mail address:* rslepaczuk@wne.uw.edu.pl (R. Ślepaczuk).

https://doi.org/10.1016/j.ribaf.2023.102052
Received 11 April 2022; Received in revised form 5 June 2023; Accepted 28 July 2023
Available online 31 July 2023

investment portfolio in order to achieve returns that exceed those achieved in the passive strategies. The selection can be made based on the methods of fundamental analysis and technical analysis. Fundamental analysis is the classic form of assessing attractiveness of a company performed in order to assess the potential future value of its shares (González-Sánchez, 2022). Among others it consists of the analysis of the company's historical and current financial statements, its long-term development strategy, the situation of the industry and economy in which the company operates analyzed in the context of their development potential and possible threats, and analysis of possible impacts on the company's results caused by the geopolitical situation.

The second family of stock selection methods is technical and statistical analysis, the main assumption of which is the possibility of forecasting the behavior of the stock price based on its historical changes (Goutte et al., 2023; Kość et al., 2019). Recurring patterns can be observed in the price charts, the ongoing identification of which is the basis for making investment decisions (Beechey et al., 2000; Anghel, 2021). On the other hand, there are numerous approaches focused on identifying the hedging properties between various assets in order to identify "safe-heaven" assets for the given financial instrument (Będowska-Sójka and Kliber, 2022). There are investors on the capital markets who use the information obtained from the fundamental and technical analysis separately, as well as concomitantly. The way to assess whether a given active strategy is in fact profitable is to compare its results with the strategy of passive capital management (Liu et al., 2023). One method of benchmarking is by relating the results to a buy-and-hold strategy on an index portfolio.

The main aim of this research is to describe, test and compare the selection of investment strategies in which investment decisions are based on signals generated by technical analysis indicators and machine learning techniques. Technical indicators such as Simple Moving Average (SMA), Moving Average Convergence Divergence (MACD), Stochastic Oscillator (SO), Relative Strength Index (RSI), and Williams' Percent Range (WPR) served as inputs to the machine learning models. The main goal of the research is to choose the best-performing strategy among the strategies constructed using various machine learning techniques such as Neural Networks, K Nearest Neighbor, Regression Trees, Random Forests, Naïve Bayes classifiers, Bayesian Generalized Linear Models, and Support Vector Machines in both Linear and Polynomial form. The second goal of the research is to compare the strategies with the method of passive capital management based on the buy-and-hold mechanism. The strategies were compared using risk and return measures, such as the annualized rate of return, the standard deviation of returns, the maximum drawdown, the Sharpe Ratio, and the Information Ratio.

The main contribution of this paper is the extension of the current achievements of scientific research by employing a wide range of machine learning methods and technical analysis to construct quantitative investment strategies and then check their robustness. The profitability of these strategies has been examined on the stock market indices of Poland (WIG20), two highly developed countries: Germany (DAX) and the USA (S&P500), as well as six countries from Central and Eastern Europe: Hungary (BUX), Czech Republic (PX), Bulgaria (SOFIX), Latvia (OMXR), Estonia (OMXT) and Lithuania (OMXV). Data used in the research consisted of High, Low, and Close daily prices of the indices in the period from 2002 to 2023, thus the scope of the research covers the periods of the great financial crisis of 2007–2009 and the COVID-19 pandemic crisis.

The originality of this study can be summarized in the below-mentioned points: the use of a wide set of ML models responsible for forecasts employed in buy-sell signals of investment strategies, the robust walk-forward procedure for testing the models, the results for a broad set of equity indices from developed and emerging markets, the quality of estimation was evaluated on independent subsets with the use of sensitivity analysis, and a more comprehensive rank approach was used to compare the final results of investment strategies.

The main research hypothesis (RH1) *states that active quantitative investment strategies based on the signals generated by machine learning models result in higher risk adjusted returns than buy-and-hold benchmark strategy*. The additional research hypotheses were as follows: (RH2) *Neural Networks generate the best (with regard to risk adjusted returns) investment signals compared to other machine learning techniques used in the research*; (RH3) *the very same machine learning strategy is considered best performing for all analyzed stock market indices*; (RH4) *returns obtained from signals generated by machine learning techniques are resistant to: changes in hyperparameters underlying the models, changes in parameters underlying the technical analysis indicators, changes in sampling and differing market conditions*. Based on the conclusions from the leading article developed by Dash and Dash (2016) which induced this research, the intuition behind the results was that machine learning based strategies can generate returns above benchmark with lower amount of investment risk. Moreover, Neural Networks strategies were suspected of being the best performing for every analyzed stock index and in every sensitivity analysis scenario.

All calculations underlying the results presented in this research were produced in R statistical software. Each scenario (base scenario and scenarios representing sensitivity analyses) took ca. 5h of computation on 2.8 GHz processor with 2 cores.

The structure of this research is composed of four chapters: the first chapter describes an overview of existing research in the field, the second chapter presents a description of the datasets, the third chapter describes the research methodology, and the fourth chapter contains results of the research, including the selection of the best performing strategies.

## 2. Literature overview

Technical analysis is a tool for detecting recurring patterns in the prices of financial instruments and making investment decisions based on them. It includes many different theories and approaches such as candlestick patterns, Elliot wave theory and Fibonacci retracement levels (Murphy, 1986). All these tools are based on separate assumptions and provide different output information. However, most of them have one feature in common - they are aimed at identifying a trend prevailing on the market and forecasting further price behavior. The vast majority of professional investors use technical analysis tools of their choice, but in the academic world it is considered unrelated to science, as it does not result directly from economic theory (Tian et al., 2002). A much more popular method among scientists is analysis of the macroeconomic environment known as fundamental analysis (Tian et al., 2002). In practice, investors combine these two methods, thus obtaining a whole picture of the market situation.

Fama (1970) formulated his efficient market hypothesis, claiming that prices of securities reflect all available information about them. In his work, he distinguished three types of market efficiency: weak, semi-strong and strong. The weak hypothesis states that the price of an instrument contains all information about its past, and therefore it is not possible to forecast the future direction of price changes on the basis of historical data. The conclusion from this hypothesis is unequivocal - technical analysis is not able to bring benefits to investors.

Despite the situation in academia at the time, Brock et al. (1922), in their empirical study on the US stock index, showed positive financial results coming from a set of three strategies based on technical analysis indicators. Their work started a new wave of scientific research aimed to determine the effectiveness of price forecasting based on historical data.

The subject of effectiveness of investment strategies based on technical analysis and machine learning techniques has been widely analyzed in scientific research. As part of the literature review, several papers fundamental to this research were discussed in the following paragraphs. The foremost paper underlying this research was an article by Dash and Dash (2016) in which the authors discussed the profitability of investment strategy constructed by the computational efficient functional link artificial neural network (CEFLANN) with Extreme Learning Machine (ELM) learning approach and technical analysis indicators such as SMA, MACD, SO, RSI and WPR. Data used in the research consisted of daily quotes of two indices: BSE SENSEX and S&P 500 from 2010 to 2014 period. Achieved returns were compared with those generated by alternative models such as Support Vector Machines, Naïve Bayesian model, K Nearest Neighbor model and Decision Tree. Results showed that CEFLANN model produced the highest returns compared to the other models.

In the paper by Jiang et al. (2012) Support Vector Machine, Multiple Additive Regression Trees, linear regression and generalized linear model (GLM) were used to solve the problem of trend prediction on NASDAQ, DJIA and S&P 500 US stock indices' daily prices. Results showed a relatively high accuracy of trend prediction achieved by the utilized techniques. Akyildirim et al. (2021) compared the efficiency of six machine learning algorithms (MLA) on high-frequency data for Bitcoin during and after the COVID-19 crisis. They showed that MLA methods outperformed the random walk and ARIMA forecasts in Bitcoin futures markets.

Huang et al. (2005) analyzed the predictive ability of Support Vector Machine, Linear Discriminant Analysis, Quadratic Discriminant Analysis and Elman Backpropagation Neural Networks. Models were applied to weekly NIKKEI 225 stock index data and incorporated several macroeconomic variables as model inputs. A model combining predictions from all of the analyzed techniques brought the best results. A machine learning approach was also used in forecasting returns of other types of assets, for example, the housing market in China (Cepni et al., 2022a,2022b). The authors revealed that after letting the machine learning models choose from all key control variables and the aligned sentiment index, the forecasting accuracy is improved at all forecasting horizons.

In the article written by Gerlein et al. (2016), six machine learning based models, including the Naïve Bayes classifier, were used to produce profitable quantitative strategies on the USDJPY, EURUSD, and EURGBP currency pairs. Models for generate positive cumulative returns in several setups.

Madan et al. (2015) applied Generalized Linear Model, Support Vector Machine and Random Forest techniques to predict the Bitcoin price change in daily as well as high frequency intervals. Authors focused on models' accuracy measurement which was relatively high for daily price change prediction in case of GLM and Random Forest models. In an article by Chen et al. (2006), authors discussed the application of Support Vector Machines and Back Propagation Neural Networks on daily close prices of six Asian stock indices: Nikkei 225, All Ordinaries, Hang Seng, Straits Times, Taiwan Weighted and KOSPI. Results showed that the analyzed models behaved better than benchmark with regard to predicted price deviation measures.

Leigh et al. (2002) described a novel approach to technical analysis bull-flag pattern recognition aiming to predict price changes. Technical indicators served as inputs to Neural Network model which was then altered with genetic algorithm in order to improve the model's coefficient of determination. Techniques were applied on New York Stock Exchange Composite Index. Calculated returns indicated the superiority of analyzed methods compared to buy-and-hold benchmark strategy.

In a paper by Lin et al. (2006), authors investigated the performance of decision trees deployed on the 'electronic stocks' of Taiwan stock market and 'technology stocks' of NASDAQ market. Predictions yielded positive returns in case of both indices.

Colianni et al. (2015) discussed construction of trading strategies based on qualitative data concerning Bitcoin observed on the Twitter portal. Linear Regression models, Support Vector Machines as well as Bernoulli and Multinomial Naïve Bayes classifiers were used. Bernoulli Naïve Bayes classifier achieved the highest accuracy in the text classification approach while Linear Regression resulted in the highest accuracy in the sentiment analysis approach compared to the remaining techniques. Kaczmarek et al. (2022) use RNN in order to determine which safe haven assets should be used when improving out-of-sample portfolio performance. Among considered assets, only long-term Treasury bonds act as a safe haven and improve the strategy performance. Other considered assets have no such potential.

One of the newest research papers (Kijewski and Ślepaczuk, 2020) in the field was testing the efficiency of ML techniques compared the performance of classical techniques with LSTM model for S&P500 index using daily frequencies in the last 20 years. They showed that the combination of strategies (ML and classical techniques) outperformed the market significantly and that the combination of signals gave the best results by diversifying the risk of single strategy mistake. Finally, they showed that LSTM with selected hyperparameters outperformed ARIMA model but at the same time LSTM model results were not robust to initial hyperparameters' assumptions. On the other hand, Zenkova and Ślepaczuk (2018) investigated the profitability of an algorithmic trading strategy based on training SVM model to identify cryptocurrencies with high or low predicted returns. They showed that equally weighted portfolio outperforms all benchmark strategies and the SVM strategy. More recent paper of Fiszeder and Orzeszko (2021) shows that the covariance matrix forecasts calculated using dynamic modelling and forecasting covariance matrices based on support vector regression using the Cholesky decomposition are more accurate than the forecasts from the benchmark dynamic conditional correlation model.

Machine learning models were also used to forecast volatility in commodity markets (Bonato et al., 2022). After estimating an extended heterogeneous autoregressive (HAR) model by means of random forests on high-frequency data covering the period from 2009 to 2020, they documented that El Niño and La Niña weather episodes increase forecast accuracy, especially at longer forecast horizons, for several of the agricultural commodities that were studied in their research. A similar attempt to employ machine learning models in the forecasting of volatility in commodity markets was made by Cepni et al. (2022a),(2022b), who found improvements in forecast accuracy, especially when they studied intermediate and long forecast horizons. Their finding was robust to various changes in the model configuration (realized variance vs. realized volatility, sample period, recursive vs. rolling-estimation window, loss function of forecast consumers).

The most recent paper of Michańków et al. (2022) investigated the efficiency of LSTM based algorithmic investment strategies test jointly on BTC and S&P500 index on 3 various frequencies (daily, hourly and 15 min). After the comprehensive sensitivity analysis, they showed the superiority of ensemble investment models built from various assets and frequencies. Ren et al. (2022) undertaken an attempt of a systematic literature review on the application of machine learning methods in cryptocurrency research is conducted. They noticed that the application of machine learning for cryptocurrencies research is increasing year over year; but at the same time, they indicated that such issues as overfitting and interpretability still persist with machine learning methods.

The aforementioned studies present a scientifically and financially interesting problem of the effectiveness of technical analysis and machine learning techniques in the development of investment strategies. An interesting question is how the strategies constructed by the researchers would behave on stock markets in less developed countries and what is the impact of the great financial crisis of 2007–2009 as well as COVID-19 pandemic crisis on the final result.

In this research, the objective is to investigate predictive ability of a set of machine learning techniques proposed by Dash and Dash (2016), namely Neural Networks, K Nearest Neighbor, Naïve Bayes, Regression Tree, Support Vector Machines and additionally Random Forest and Bayesian Generalized Linear Model discussed in other papers. Technical analysis indicators proposed by Dash and Dash (2016) were used as inputs to aforementioned machine learning models. Analyzed set of indicators consisted of Simple Moving Average (SMA), Moving Average Convergence Divergence (MACD), Stochastic Oscillator (STOCH), Relative Strength Index (RSI) and Williams' Percent Range (WPR) with the underlying parameters similar to those proposed in the paper.

## 3. Data

### 3.1. Data description

Data used in the research was downloaded from the website *http://www.stooq.pl/* and contains information about HLC (High Low Close) prices of selected stock indices. For the analysis, authors chose the stock market index of their domestic market – WIG20 (Poland), two highly liquid global indices from developed countries – DAX (Germany) and S&P500 (USA) as well as a group of less liquid Central and Eastern European countries' indices: BUX (Hungary), PX (Czech Republic), SOFIX (Bulgaria), OMXR (Latvia), OMXT (Estonia) and OMXV (Lithuania). Data for each index contained quotes from 2002 to 01–01 to 2023–03–31. Due to suspension of quotation observed for certain indices as well as differing holiday calendars, there were dates with no quotes available. This limitation was remediated by omitting those dates in the analysis resulting in a number of observations differing among indices.

### 3.2. Sampling

Research employed dynamic estimation windows which means that the underlying parameters of the models were periodically recalibrated to reflect current market behaviors. Observations from the beginning of the available period were initially trimmed in order for the overall number of observations for each index to be easily divisible into equal subsets. Calibration of models' parameters was conducted on 200 trading day window (in-sample) and then model predictions were applied onto next 20 trading day window (out-of-sample). For each subsequent dynamic window iteration, in-sample and out-of-sample moved by 20 trading days. The process

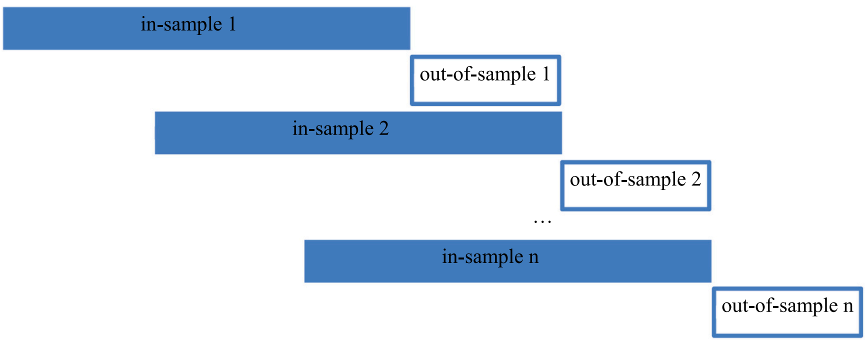


**Fig. 2.1.** Sampling process overview. Note: Figure illustrates the sampling process showing how in-sample and out-of-sample subsets are derived from the overall dataset.

is shown on Fig. 2.1.

Table 2.1 presents the number of observations after aforementioned trimming process for each of analyzed indices with the corresponding number of in-sample and out-of-sample subsets created as well as the start and end date of the overall sample period.

### 3.3. Initial data analysis

Initial data analysis was conducted in order to assess the distribution of target variable (discrete returns derived from Close prices - target variable in the models described in detail at the beginning of the next section) used for modeling purposes.

Table 2.2 presents mean, minimum, 25th percentile, 50th percentile (median), 75th percentile and maximum values of the variable for each of the analyzed indices. No outliers or data quality issues were identified.

## 4. Research methodology

### 4.1. General model formula and target variable

One of the general goals of the quantitative investment strategies development is to construct models with price predictive ability. Machine learning models developed in this research belong to the family of supervised models which means that they are fed with pairs of input (technical analysis indicators) and target (stock indices returns) variables. Information coming from the input and target pairs was used to calibrate each models' coefficients in each of the in-sample periods. Those coefficients were then applied to the inputs in the following out-of-sample periods in order to predict the target variable in those periods. The general modeling matter can be described using Formula (1).

$$f(Y) = f(X) + \varepsilon \tag{1}$$

where: Y – vector of target variable (returns),

X – matrix of independent variables (set of technical indicators),

f(.) – function applied to transform the variables (depending on the model),

ε – vector of random errors.

Target (dependent) variable in this research is defined as a discrete return on the asset calculated from the observed Close prices as described by Formula (2).

$$r_t = \frac{C_t - C_{t-1}}{C_{t-1}} \tag{2}$$

where: $r_t$ – discrete return in period t,

$C_t$ – close price in period t.

### 4.2. Technical analysis indicators

Research focuses on strategies based on a set of 5 technical analysis indicators: Simple Moving Average (SMA), Moving Average Convergence Divergence (MACD), Stochastic Oscillator (STOCH), Relative Strength Index (RSI) and Williams' Percent Range (WPR) as proposed by Dash and Dash (2016). Technical indicators were then used as an input to machine learning models. The following section describes formulas used for calculation of each analyzed indicators.

#### 4.2.1. Simple Moving Average (SMA)

Simple Moving Average is an average price of an instrument calculated on historical observations up to the reference date as proposed by Feller (1991). Base level analyzed for parameter representing the number of periods n = 15 as proposed by Dash and Dash

**Table 2.1**
Data sampling overview.

| Index | Observations | Subsets | Start date | End date |
|---|---|---|---|---|
| WIG20 | 5280 | 254 | 2002–03–04 | 2023–03–31 |
| DAX | 5360 | 258 | 2002–02–27 | 2023–03–31 |
| SPX | 5280 | 254 | 2002–04–12 | 2023–03–31 |
| BUX | 5240 | 252 | 2002–04–17 | 2023–03–31 |
| PX | 5280 | 254 | 2002–03–18 | 2023–03–31 |
| SOFIX | 5200 | 250 | 2002–02–27 | 2023–03–31 |
| OMXR | 5280 | 254 | 2002–03–06 | 2023–03–31 |
| OMXT | 5280 | 254 | 2002–04–03 | 2023–03–31 |
| OMXV | 5240 | 252 | 2002–04–05 | 2023–03–31 |

Note: Table presents descriptive statistics for each stock index: number of observations, number of subsets, start date and end date of observation period in the dataset.

**Table 2.2**
Descriptive statistics of all analyzed indices.

| Measure / Index | WIG20 | DAX | SPX | BUX | PX | SOFIX | OMXR | OMXT | OMXV |
|---|---|---|---|---|---|---|---|---|---|
| Mean | 0.0001 | 0.0003 | 0.0003 | 0.0004 | 0.0003 | 0.0004 | 0.0004 | 0.0005 | 0.0005 |
| Minimum | -0.1328 | -0.1224 | -0.1198 | -0.1188 | -0.1494 | -0.1074 | -0.1507 | -0.1006 | -0.1125 |
| 25th percentile | -0.0075 | -0.0061 | -0.0045 | -0.0072 | -0.0053 | -0.0043 | -0.0046 | -0.0033 | -0.0027 |
| 50th percentile | 0.0002 | 0.0008 | 0.0007 | 0.0006 | 0.0007 | 0.0003 | 0.0002 | 0.0006 | 0.0005 |
| 75th percentile | 0.0078 | 0.0072 | 0.0058 | 0.0080 | 0.0066 | 0.0049 | 0.0053 | 0.0044 | 0.0040 |
| Maximum | 0.0850 | 0.1140 | 0.1158 | 0.1408 | 0.1316 | 0.0875 | 0.1285 | 0.1286 | 0.1163 |

Note: Table presents descriptive statistics for returns of all analyzed indices: mean, minimum, 25th percentile, 50th percentile (median), 75th percentile and maximum.

(2016) and for the purpose of sensitivity analysis, n = {14; 16} were used in order to verify robustness of the models. Input used later in the models is derived with the usage of Formula (3). It is a measure of how distant the current price is from its SMA:

$$SMA_{signal} = P_t - SMA \tag{3}$$

#### 4.2.2. Moving Average Convergence Divergence (MACD)

MACD is an indicator developed by Appel (2005) which incorporates several Exponential Moving Averages (EMA) into its derivation. EMA is described as a moving average which assigns exponentially decreasing weights (older the observation, lower the weight) to each of the historical observations. MACD indicator is composed of 2 distinct time series: MACD line and signal line. MACD line is defined as the difference between the long EMA and short EMA with length of the periods set on the level of n = 26 and n = 12 periods. Signal line is defined as EMA with parameter n = 9. Above mentioned parameters are considered base parameters in this research. For the purpose of sensitivity analysis, n = {25; 27} for long EMA, n = {11; 13} for short EMA and n = {8; 10} for signal EMA were used in order to verify robustness of the models.

Input used later in the models is derived with the usage of Formula (4). It is a measure of how distant the MACD line is from the signal line:

$$MACD_{signal} = MACD \quad line - signal \quad line \tag{4}$$

#### 4.2.3. Stochastic Oscillator (STOCH)

Invented by Lane (1984), Stochastic Oscillator incorporates HLC (High Low Close) data into its calculation. STOCH comprises of 3 time series: fast %K, fast %D and slow %D. %D andLane proposed that fast %K should be calculated with parameter n = 14 while fast %D and slow %D as SMAs with n = 3 periods, those levels are treated as base levels in this research. Sensitivity analysis was conducted using n = {13; 15} for fast %K and n = {2; 4} for fast %D and slow %D.

#### 4.2.4. Relative Strength Index (RSI)

Base level parameter n = 14 was employed as proposed by Wilder (1978). In sensitivity analysis, n = {13;15} were used in order to verify robustness of the models.

#### 4.2.5. Williams' percent range (WPR)

Developed by Williams (1979), WPR is a form of a price oscillator using HLC (High Low Close) data. It is calculated similarly to fast %K in Stochastic Oscillator. Similarly to RSI indicator, base level is analyzed for parameter n = 14 as proposed by the author of WPR while sensitivity analysis is conducted using n = {13; 15}.

All analyzed technical indicators are lagged by one period before being used as predictors for returns in the models in order to avoid the so-called look ahead bias involving making decisions in the same period for which the given signal was generated.

### 4.3. Machine learning techniques

Research analyzed eight supervised machine learning models with the majority of them proposed in the paper by Dash and Dash (2016) and others discussed in the remaining papers. Employed techniques included Neural Networks, K Nearest Neighbor, Random Forest, Regression Tree, Naïve Bayes, Bayesian Generalized Linear Model and Support Vector Machines in both Linear and Polynomial form. Following subsections describe each of the models and discuss the hyperparameters used to conduct one of the sensitivity analysis exercises.

#### 4.3.1. Neural Networks (NN)

Models and corresponding strategies referred in this research as Neural Networks were developed using the Extreme Learning Machine (ELM) approach as proposed by Dash and Dash (2016). ELM in the configuration applied is a feedforward neural network with only one hidden layer which is the reason why this approach is computationally efficient compared to other neural network related techniques. The number of neurons in the input layer is equal to the number of input technical analysis indicators. In each of the in-sample estimations, a model is trained using a number of neurons in hidden layer varying from 1 to twice the size of input layer and the best performing in-sample variant is then chosen. The activation function in a tansig (tangent-sigmoid transfer function) form producing continuous values in the range from -1 to 1 (intuitive for return prediction) was applied to compute the output trading signal in the output layer consisting of one neuron. The model was implemented in the form discussed by Huang et al. (2006). Output from Extreme Learning Machine algorithm has a form described by Formula (5).

$$f_n(X) = \sum_{i=1}^{n} \beta_i h_i(X) \tag{5}$$

where: $f_n$ – predicted output from the model,

$X$ – matrix of independent variables,

$n$ – number of neurons in hidden layer,

$\beta_i$ – weight of hidden neuron i,

$h_i$ – output of hidden neuron i.

The base activation function analyzed was the tansig transformation and, for the purpose of sensitivity analysis, two alternative functions producing outputs in the same range were chosen: sin (sine transfer function) and satlins (symmetric saturating linear transfer function).

#### *4.3.2. K nearest neighbor (KNN)*

Research employs K Nearest Neighbor model in its regression version. The output prediction of the model is the average value of observed target variable for k nearest neighbors identified based on the levels of input independent variables. Models were implemented in a form proposed by Altman (1991). KNN algorithm is described with Formula (6).

$$p = \frac{\sum_{i=1}^{k} y_i}{k} \tag{6}$$

where: $p$ – predicted output from the model,

$y_i$ – observed target variable for nearest neighbor observation i,

$k$ – number of nearest neighbors included in the calculation.

Identification of k nearest neighbors is based on determination of high dimensional Euclidean distance between independent variables of analyzed observations as described in Formula (7).

$$d_{i,j}^2 = \sum_{l=1}^{n} (x_{l,i} - x_{l,j})^2 \tag{7}$$

where: $d_{i,j}^2$ – high dimensional Euclidean distance between observations i and j,

$x_{l,i}$ – independent variable l for observation i,

$n$ – number of independent variables.

The hyperparameter chosen for sensitivity analysis was an optimization metric with the Root Mean Square Error (RMSE) as the base metric and alternative metrics being the coefficient of determination (Rsquared) and Mean Absolute Error (MAE).

#### *4.3.3. Random Forest (RF)*

Random Forest model in a regression form used in this research is a statistical modelling framework consisting of random generation of multiple decision trees with each of the trees producing a distinct prediction for target variable. Those predictions are then averaged to calculate the final output. Models were implemented in a form discussed by Breiman (2001). Process of the output generation is described by Formula (8).

$$p = \frac{\sum_{j=1}^{m} \sum_{i=1}^{n} W(x_i, x') y_i}{m} \tag{8}$$

where: $p$ – predicted output from the model,

$y_i$ – observed target variable for observation i,

$x_i$ – vector of independent variables for observation i,

$x'$ – vector of independent variables for observation in testing sample,

$W(x_i, x')$ – weight function of $x_i$ relative to $x'$,

$n$ – number of observations in training sample,

$m$ – number of generated random trees.

The hyperparameter chosen for sensitivity analysis was an optimization metric with the RMSE as the base metric and alternative metrics being the Rsquared and MAE.

#### *4.3.4. Regression Tree (RT)*

Recursive partitioning Regression Trees are a version of Decision Trees from the Classification and Regression Trees (CART) family with continuous target variable. Data is split in recursive manner in order to generate optimal decision algorithm for target variable prediction. Model inputs (independent variables) are reflected in the tree branches from which, after a set of recursive partitioning, final leaves with the computed target variable are produced. Models were implemented in a form proposed by Breiman et al. (1984).

Fundamental algorithms of Regression Tree are similar to those of Random Forest model. The hyperparameter chosen for sensitivity analysis was an optimization metric with the RMSE as the base metric and alternative metrics being Rsquared and MAE.

#### 4.3.5. Naïve Bayes (NB)

Naïve Bayes is a probabilistic classifier incorporating the assumption of naïve independence between input variables. As a classification method, it produces binary outputs (classes) computed from conditional posteriori probabilities. Research used {− 1;1} classes representing buy and sell trading signals thus the Naïve Bayes model was the single model implemented without an additional 'neutral' signal. The model was implemented in a form proposed by Murty and Devi (2011). Output from the model is described by Formula (9).

$$p = q(C_k)\prod\nolimits_{i=1}^{n} q(x_i|C_k) \tag{9}$$

where: p – predicted output from the model,

$q(.)$ – probability of (.),

$C_k$ – class k,

$x_i$ – independent variable i,

n – number of independent variables.

The hyperparameter chosen for sensitivity analysis was an optimization metric with Accuracy (number of correct predictions divided by the total number of predictions) as the base metric and one alternative metric being the Cohen's kappa (Kappa).

#### 4.3.6. Bayesian Generalized Linear Model (BGLM)

Generalized Linear Model is a generalization of linear regression models which among others allows for target variable transformations via a link function e.g. logit which was used in this research. Target variable prediction is computed as a linear combination of input variables. BGLM uses the Bayesian approach to model fitting instead of the Frequentist approach. A priori distributions of inputs and the likelihood function are used for a posteriori estimation of model parameters. Models were implemented in forms discussed by Nelder and Wedderburn (1972) and Dempster et al. (1977). Model output is described by Formula (10).

$$p = l^{-1}(X\beta) \tag{10}$$

where: p – predicted output from the model,

l – link function,

X – matrix of independent variables,

β – vector of coefficients fitted by the model using the Bayesian approach.

The hyperparameter chosen for sensitivity analysis was an optimization metric with RMSE as the base metric and alternative metrics being Rsquared and MAE.

#### 4.3.7. Support Vector Machine Linear (SVML)

Regression version of Support Vector Machine models was analyzed in this research. SVM models generate multiple hyperplanes aiming to separate input independent variables and search for the most optimal solution allowing for the best prediction of the continuous target variable. Linear type of SVM was implemented in a form discussed by Cortes and Vapnik (1995). The objective function of the model is described by Formula (11). It has to be determined by identification of the optimal hyperplane using the minimization problem presented in Formula (12). Additionally, all of the model residuals have to fulfill the condition described by Formula (13).

$$f(X) = X^{'}\beta + b \tag{11}$$

$$\frac{1}{2}\beta^{'}\beta \tag{12}$$

$$\forall i : |y_i - (X_i^{'}\beta + b)| \leq \varepsilon \tag{13}$$

where: f(X) – function of X, target to determination,

X – matrix of independent variables,

β – vector of coefficients,

b – offset intercept,

$y_i$ – observed dependent variable for observation i,

$X_i$ – vector of independent variables for observation i,

ε – random error.

The hyperparameter chosen for sensitivity analysis was an optimization metric with RMSE as the base metric and alternative metrics being Rsquared and MAE.

#### 4.3.8. Support Vector Machine Polynomial (SVMP)

Polynomial Support Vector Machine models employed in the research are a version of SVM models incorporating the polynomial kernel function transforming model inputs and computing high-dimensional hyperplanes. Polynomial type of SVM was implemented in a form discussed by Boser et al. (1992). Polynomial kernel function has a form described by Formula (14).

$$k(x_i, x_j) = (1 + x_i^{'} x_j)^d \tag{14}$$

where: k(.) – kernel function,

$x_i$, $x_j$ – vectors of independent variables,

d – degree of the polynomial.

The hyperparameter chosen for sensitivity analysis was an optimization metric with RMSE as the base metric and alternative metrics being Rsquared and MAE.

### 4.4. Investment strategies construction

Construction of quantitative investment strategies required multiple computational steps and definition of trading rules. This section describes the process of inputting technical analysis indicators into the machine learning models as well as the process of generation of trading signals from the models' outputs.

#### 4.4.1. Extended model formula

The general model Formula (1) can be further extended to Formula (15) to present each particular independent variable described in detail in the Technical Analysis Indicators section.

$$f(y) = f(SMA_{signal}) + f(MACD_{signal}) + f(fast\%K) + f(fast\%D) + f(slow\%D) + f(RSI) + f(WPR) + \varepsilon \tag{15}$$

where: y – vector of target variable (returns),

f(.) – function applied to transform the variables (depending on the model),

ε – vector of random errors.

#### 4.4.2. Model inputs transformation

Input independent variables (technical analysis indicators) were rescaled before being fed to the models. The process was conducted using a version of min-max normalization technique which produces outputs in range from -1 to 1. This technique was chosen for two reasons: it is intuitive as the machine learning models produce output variable that is also ranging from -1 to 1 and because it causes the input data to be more comparable. Process of min-max normalization (rescaled to range from -1 to 1) is described by Formula (16) as proposed by Han et al. (2011).

$$x_t^{'} = \frac{x_t - \min(x)}{(x) - \min(x)} * 2 - 1 \tag{16}$$

where: $x_t^{'}$ – transformed value of variable in period t,

$x_t$ – original value of variable in period t,

min(x)/max(x) – minimum/maximum value of the variable in all analyzed periods.

#### 4.4.3. Model outputs transformation (investment signals creation)

Machine learning models used in this research can be divided into two groups: classification models (Naïve Bayes) and regression models (remaining techniques). Outputs (returns predictions) and corresponding trading signals for each of the incorporated models constitute a distinct investment strategy. Classification models produce a binary output $\{-1;1\}$ while regression models produce continuous output ranging from -1 to 1. The question to be answered is how to translate the outputs into trading signals.

For classification models a simplistic approach was undertaken: $-1$ output translates to sell signal while $+1$ output translates to buy signal. Due to the binary output, this technique does not allow to produce neutral investment signals.

In the case of regression models, the outputs were highly dispersed and non-comparable among the models in the sense of distribution measures therefore not allowing to set a fixed signal thresholds based on absolute values of the outputs. The decision was made that the most universal approach to signal generation will be to calculate quantiles of the output distributions for each of the analyzed models. 40th quantile and 60th quantile were applied as the thresholds for buy, sell and neutral signals. Signal $+1$ translates to buy signal, $-1$ to sell signal and 0 to neutral signal. The process of signal generation for regression models is described by Formula (17).

$$signal_t = \{1 y_t \geq q_{0.6} 0 q_{0.4} < y_t < q_{0.6} - 1 y_t \leq q_{0.4} \tag{17}$$

where: $signal_t$ – trading signal in period t,

$y_t$ – output (prediction) generated by the model for period t,

$q_\alpha$ – quantile of particular out-of-sample outputs corresponding to probability α.

The process of entering a financial position was based on buy, sell and neutral signals. Neutral signal is interpreted as not taking a position or exiting an existing one. To calculate the return from a given strategy for each date, signal was multiplied by the observed discrete return of a given financial instrument which is described by Formula (18).

$$r_t^{strategy} = r_t^{index} * signal_{t-1}^{strategy} \tag{18}$$

where: $r_t^{strategy}$ – discrete return from the strategy in period t,

$r_t^{index}$ – discrete return from stock index (financial instrument analyzed) in period t,

$signal_t^{strategy}$ – signal generated by the strategy in period t.

Returns from the strategies were aggregated for every out-of-sample period in order to produce return time series and compare the strategies among each other and additionally with the benchmark strategy. Research investigates which of the models (strategies) gives the most desired results. The effectiveness of the quantitative strategies was compared to that of a buy-and-hold strategy which is based on a market portfolio (benchmark). The buy-and-hold strategy involves buying an instrument at the beginning of the period under analysis and selling it at the end of the period, so it can be interpreted as an absolute measure of market movements.

### 4.5. Risk and return measures

Research incorporates a wide range of performance indicators used to assess the quality of developed investment strategies. In order to appropriately compare the strategies, not only the accumulated profits but also the risks should be considered. Measures and ratios used in the analysis included the compound annual growth rate, standard deviation of returns, maximum capital drawdown, Sharpe Ratio and Information Ratio.

#### 4.5.1. Compound Annual Growth Rate (CAGR)

The rate of return is the most frequently used measure in portfolio efficiency studies. It is a measure illustrating how much on average capital has grown in each year of investment. To calculate it, the following Formula (19) is used (Anson et al., 2010):

$$R = CAGR(t_0, t_n) = \left(\frac{V(t_n)}{V(t_0)}\right)^{\frac{1}{t_n - t_0}} - 1 \tag{19}$$

where: $CAGR(t_0,t_n)$ – Compound Annual Growth Rate,

$V(t_0)$ – initial value of an investment,

$V(t_n)$ – closing value of an investment,

$t_0$ – calculation start year,

$t_n$ – calculation end year.

#### 4.5.2. Adjusted Sharpe Ratio (SR)

The rate of return itself does not contain any information about risk. The solution to this problem was proposed by Sharpe (1966) who introduced Sharpe's coefficient. This research used its simplified version, which does not contain information about the risk-free rate. It is calculated by dividing the annualized rate of return by the annualized standard deviation of rates of return in a given period. The standard deviation illustrates volatility of returns and is considered as a risk measure in which greater volatility indicates a higher investment risk. Considering the aforementioned information, when comparing strategies, the better performing one is the one with the higher Adjusted Sharpe Ratio. For the purpose of this research, the measure was floored at 0 as the negative values are often deemed meaningless in the scientific world. It is calculated with Formula (20).

$$SR = \max\left\{\frac{R}{\sigma R}; 0\right\} \tag{20}$$

where: SR – Adjusted Sharpe Ratio,

R – annualized rate of return,

$\sigma R$ – annualized standard deviation of returns.

#### 4.5.3. Maximum drawdown (MDD)

The maximum drawdown represents the maximum decrease in accumulated capital over the entire investment horizon. When analyzing the rate of return, it is worth investigating whether the portfolio has not recorded significant drops in value in the analyzed horizon, which would indicate its instability. The most frequently used measure for this purpose is the maximum drawdown which describes that risk. It is a difference between the value of capital at the lowest point and the value at the previous highest peak divided by the value at that peak. The final value is usually shown as a percentage. In this research, the measure is always presented as positive value, so in superior investment strategies the maximum decline should be as low as possible. This situation will represent a lower risk of a managed portfolio. Measure is calculated as follows (Magdon-Ismail et al., 2004) using Formula (21):

$$MDD = -\frac{TMin - PMax}{PMax} \tag{21}$$

where: MDD – maximum drawdown,
TMin – minimum price level,
PMax – previous maximum price level.

#### 4.5.4. Calmar Ratio (CR)

The Calmar Ratio is a very useful extension of the maximum drawdown described before. It is another risk adjusted return measure that results from dividing the annualized rate of return by the maximum drawdown expressed in absolute value. Due to its structure, it is a measure that has a priority in application before the maximum drawdown (Bacon, 2021). As in the case of the Sharpe Ratio, the better performing strategies are those with the higher value of the Calmar Ratio. The applied formula (Young, 1991) is described by Eq. (22). For the purpose of this research, the measure was floored at 0 as the negative values are often deemed meaningless in the scientific world.

$$CR = \frac{\max\{R; 0\}}{MDD} \tag{22}$$

where: R – annualized rate of return,
MDD – maximum drawdown.

#### 4.5.5. Sortino Ratio (SoR)

The Sortino Ratio is a modification of Sharpe Ratio described before. In a general form, in a numerator it calculates a difference between annualized rate of return and a custom required rate of return or target return (which in this paper is assumed to be 0) and divides it by standard deviation of downside moves. This means that Sortino Ratio penalizes only for downside volatility. Similar to Sharpe Ratio, the better performing strategies are those with the higher value of the Sortino Ratio. The applied formula (Sortino and Price, 1994) is described by Eq. (23). For the purpose of this research, the measure was floored at 0 as the negative values are often deemed meaningless in the scientific world.

$$SoR = \frac{\max\{R; 0\}}{\sigma \mathrm{Rd}} \tag{23}$$

where: SoR – Sortino Ratio,
R – annualized rate of return,
σRd – annualized downside deviation of returns.

#### 4.5.6. Information Ratio* (IR*)

For the purpose of this research, an adjusted Information Ratio definition was introduced described in Formula (24) as proposed by Kość et al. (2019). The measure was floored at 0 as the negative values are often deemed meaningless in the scientific world.

$$IR^* = SR * CR = \frac{(\max\{R; 0\})^2}{\sigma R * MDD} \tag{24}$$

where: IR* – Information Ratio* ,
R – annualized rate of return,
σR – annualized standard deviation of returns,
MDD – maximum drawdown.

There are many measures of risk and return used by researchers to compare the performance of investment strategies, but each of them is suitable for analyzing different types of instruments contained in a portfolio (Bacon, 2021). The aforementioned measures and ratios fully cover the needs of this research and will allow for an objective assessment of the quality of discussed strategies.

## 5. Empirical research

### 5.1. Identification of the best performing strategy

In order to determine the best performing strategy for a given stock index, the out-of-sample results produced by eight analyzed machine learning models were translated into trading signals and compared using risk and return measures. Trading signals produced by each machine learning model constitute a separate investment strategy. This section describes the results obtained by the strategies for each analyzed stock index. Results for the most liquid indices in the analyzed group i.e. WIG20 (Poland), DAX (Germany) and S&P500 (USA) are described in detail whereas the results from remaining CEE stock indices are presented in an aggregated approach. The evaluation of the obtained results was based on a graphical analysis i.e. equity lines (presenting how 1$ of initial investment would row over the analyzed time period), daily returns and drawdown lines as well as the calculated values of risk and return measures. The main measure selected for strategy comparison was IR* as it contains the highest amount of information about the performance (it combines information about returns, standard deviation and maximum drawdown). Eight strategies built from eight machine learning models are also compared with the benchmark strategy i.e. buy-and-hold which in the following sections is always presented with the name of the relevant index. Results presented in this section were referred to as a base scenario in the following sensitivity analysis sections.

#### 5.1.1. Investment strategies comparison for WIG20 (Poland)

As described in the data section, in case of WIG20 index, 254 subsamples were created with dates ranging from 2002-03–04 to 2023–03–31. Out-of-sample results were aggregated into a time series of discrete returns for every analyzed strategy and compared with each other as well as with the benchmark.

Fig. 4.1 presents the equity lines, daily returns and drawdown lines for all analyzed strategies. Support Vector Machine strategies are dominant in the case of WIG20 index with its Polynomial version outperforming the rest of the strategies significantly. In most of the periods of increased index volatility, the equity lines are also increasing (e.g. in 2007–2009 great financial crisis period) while in COVID-19 pandemic crisis period this relationship is contradictory.

Another measure worth graphical analysis is the drawdown lines chart, i.e. the decline in the value of the portfolio calculated from the previous peak to the current trough achieved in the analyzed periods. The red line represents the buy-and-hold strategy which achieved the highest equity drops from all analyzed strategies in the 2007–2009 great financial crisis period. Cyan line representing Naïve Bayes results suggests that the biggest drops in COVID-19 pandemic period were observed for that strategy. Overall, Naïve Bayes, Random Forest and Regression Tree strategies obtained the worst drawdowns from the group of machine learning strategies. The remaining strategies behaved in a more efficient manner. Portfolio value drops are inevitable in active investment management, but it is always worth minimizing them.

In the further step of the analysis, risk and return measures obtained by each investment strategy were compared among each other. The summary of those measures is presented in Table 4.1. For the IR* measure, the best score of 1.72 was obtained by Polynomial SVM model whereas the result for the benchmark strategy was 0.0001. As IR* was chosen as the most decisive in this research, Polynomial SVM was considered the best performing from all analyzed strategies. In the following sections strategies will be compared for the remaining indices.

#### 5.1.2. Investment strategies comparison for DAX (Germany)

For DAX index, 258 subsamples were created with dates ranging from 2002-02–27 to 2023–03–31. Fig. 4.2 presents equity lines, daily returns and drawdown lines for all analyzed strategies. Regression Tree strategy produced equity line with the highest overall return. The next best strategies were those produced by Linear Support Vector Machine and Bayesian Generalized Linear Model.

As in case of WIG20 index, in most of the periods of increased volatility of returns, the equity lines are also increasing e.g. in 2007–2009 great financial crisis period while in COVID-19 pandemic crisis period, this relationship is contradictory.

The highest equity drops from all analyzed strategies are observed mostly in the 2007–2009 great financial crisis period and COVID-19 pandemic period. Naïve Bayes, Polynomial Support Vector Machine, K Nearest Neighbor and Neural Networks models produced drawdowns worse than the benchmark strategy.

Table 4.2 presents risk and return measures calculated for all analyzed DAX strategies. The best score of 0.34 for IR* measure was obtained by SVML model whereas the result for the benchmark strategy was 0.06, therefore SVML model was considered best performing from all analyzed strategies.

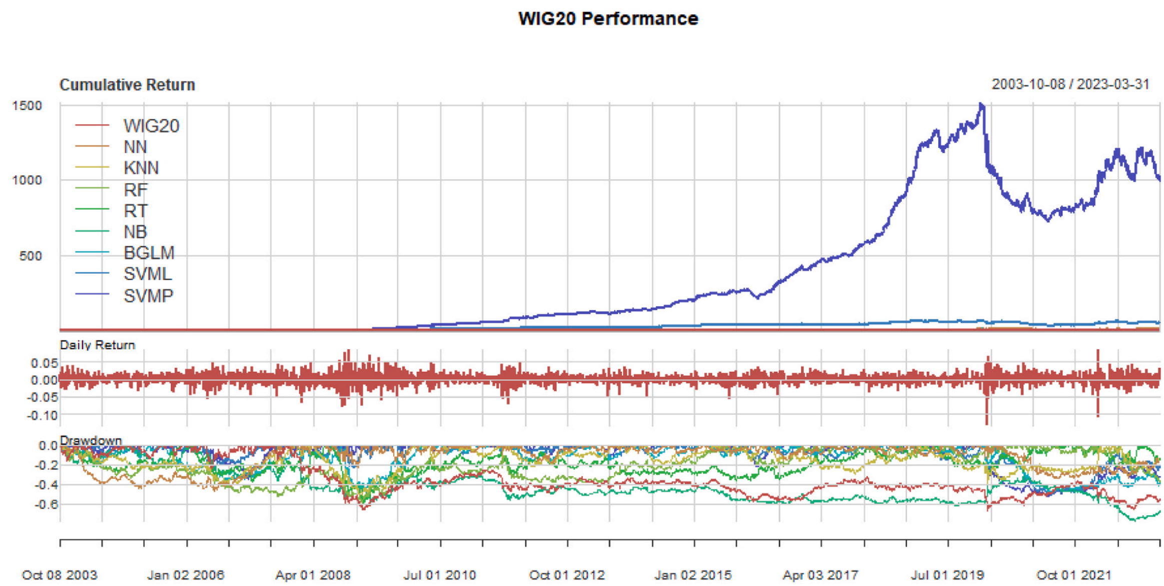


**Fig. 4.1.** Equity lines, daily returns and drawdown lines for WIG20 (Poland). Note: Figure shows equity lines, daily returns and drawdown lines for every strategy constructed on WIG20 index in the period from 2003-10–08 to 2023–03–31.

**Table 4.1**
Risk and return measures for WIG20 (Poland).

| Measure | WIG20 | NN | KNN | RF | RT | NB | BGLM | SVML | SVMP |
|---|---|---|---|---|---|---|---|---|---|
| CAGR | 0.44% | 12.90% | 4.61% | 2.87% | 2.16% | -5.15% | 9.91% | 22.55% | 42.85% |
| Annual. Std Dev | 23.04% | 20.34% | 20.74% | 20.81% | 22.93% | 21.34% | 20.70% | 20.38% | 20.57% |
| Adj Sharpe | 0.0193 | 0.6341 | 0.2224 | 0.1377 | 0.0941 | -0.2415 | 0.4787 | 1.1062 | 2.0830 |
| MDD | 66.67% | 46.56% | 51.39% | 60.49% | 58.74% | 77.15% | 48.44% | 50.83% | 51.79% |
| IR* | 0.0001 | 0.1756 | 0.0200 | 0.0065 | 0.0035 | 0.0000 | 0.0979 | 0.4907 | 1.7234 |
| SoR | 0.0119 | 0.0642 | 0.0290 | 0.0216 | 0.0186 | 0.0000 | 0.0515 | 0.1038 | 0.1801 |
| CR | 0.0067 | 0.2770 | 0.0897 | 0.0474 | 0.0367 | 0.0000 | 0.2045 | 0.4436 | 0.8274 |

Note: Table shows risk and return measures for strategies constructed on WIG20 index. The first column represents buy-and-hold strategy. Presented measures include: CAGR, annualized standard deviation, adjusted Sharpe Ratio, Maximum Drawdown, IR* , Sortino Ratio and Calmar Ratio. Bolded font indicates the best performance measure for all tested methods.

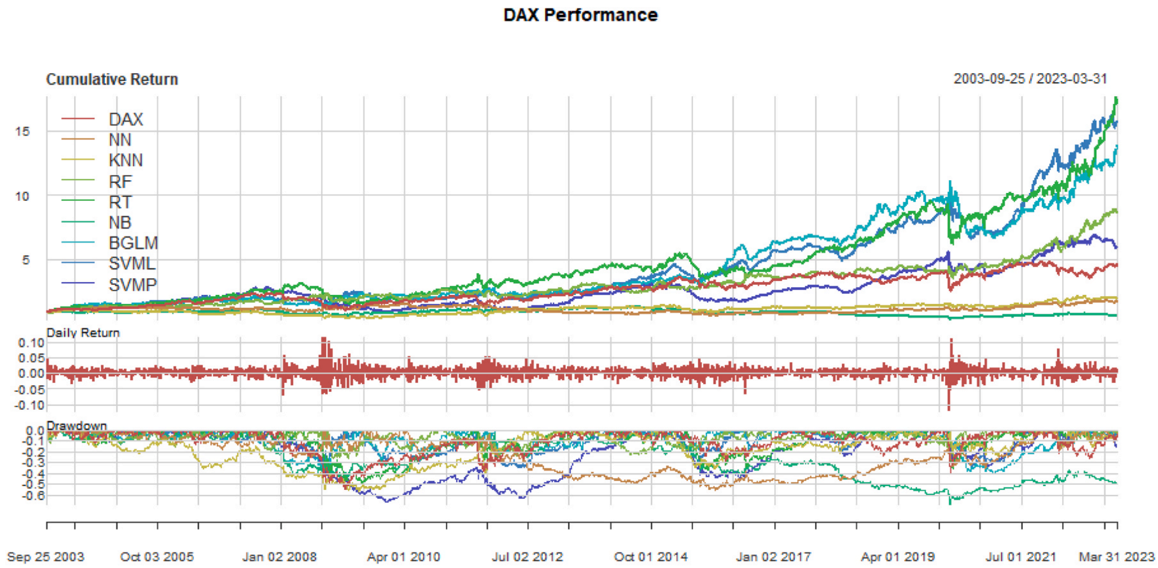


**Fig. 4.2.** Equity lines, daily returns and drawdown lines for DAX (Germany). Note: Figure shows equity lines, daily returns and drawdown lines for every strategy constructed on DAX index in the period from 2003-09–25 to 2023–03–31.

**Table 4.2**
Risk and return measures for DAX (Germany).

| Measure | DAX | NN | KNN | RF | RT | NB | BGLM | SVML | SVMP |
|---|---|---|---|---|---|---|---|---|---|
| CAGR | 8.18% | 3.18% | 3.56% | 11.69% | 15.56% | -1.78% | 14.29% | 15.01% | 9.60% |
| Annual. Std Dev | 21.08% | 18.97% | 19.12% | 19.10% | 21.01% | 18.35% | 18.67% | 18.76% | 19.09% |
| Adj Sharpe | 0.3879 | 0.1676 | 0.1864 | 0.6120 | 0.7410 | -0.0970 | 0.7651 | 0.8001 | 0.5027 |
| MDD | 54.77% | 55.82% | 56.15% | 26.52% | 50.73% | 68.14% | 40.04% | 35.10% | 66.07% |
| IR* | 0.0579 | 0.0095 | 0.0118 | 0.2697 | 0.2273 | 0.0000 | 0.2731 | 0.3421 | 0.0730 |
| SoR | 0.0426 | 0.0231 | 0.0253 | 0.0619 | 0.0718 | 0.0000 | 0.0763 | 0.0780 | 0.0502 |
| CR | 0.1493 | 0.0570 | 0.0635 | 0.4407 | 0.3068 | 0.0000 | 0.3569 | 0.4276 | 0.1452 |

Note: Table shows risk and return measures for strategies constructed on DAX index. The first column represents buy-and-hold strategy. Presented measures include: CAGR, annualized standard deviation, adjusted Sharpe Ratio, Maximum Drawdown, IR* , Sortino Ratio and Calmar Ratio. Bolded font indicates the best performance measure for all tested methods.

### *5.1.3. Investment strategies comparison for S&P500 (USA)*

In case of S&P500 index, 254 subsamples were created with dates ranging from 2002–04–12 to 2023–03–31. Fig. 4.3 presents equity lines, daily returns and drawdown lines for all analyzed strategies. Situation was similar to the one observed for WIG20, Polynomial Support Vector Machine strategy was dominant also for S&P500 index. The next best overall return was observed for Bayesian Generalized Linear Model.

As in case of WIG20 and DAX indices, in most of the periods of increased volatility of returns, the equity lines are also increasing. For S&P500 index it was observed not only in 2007–2009 great financial crisis period but also in COVID-19 pandemic crisis period. The worst capital drawdown was observed for Naïve Bayes model, while the remaining models behaved better than the benchmark in that

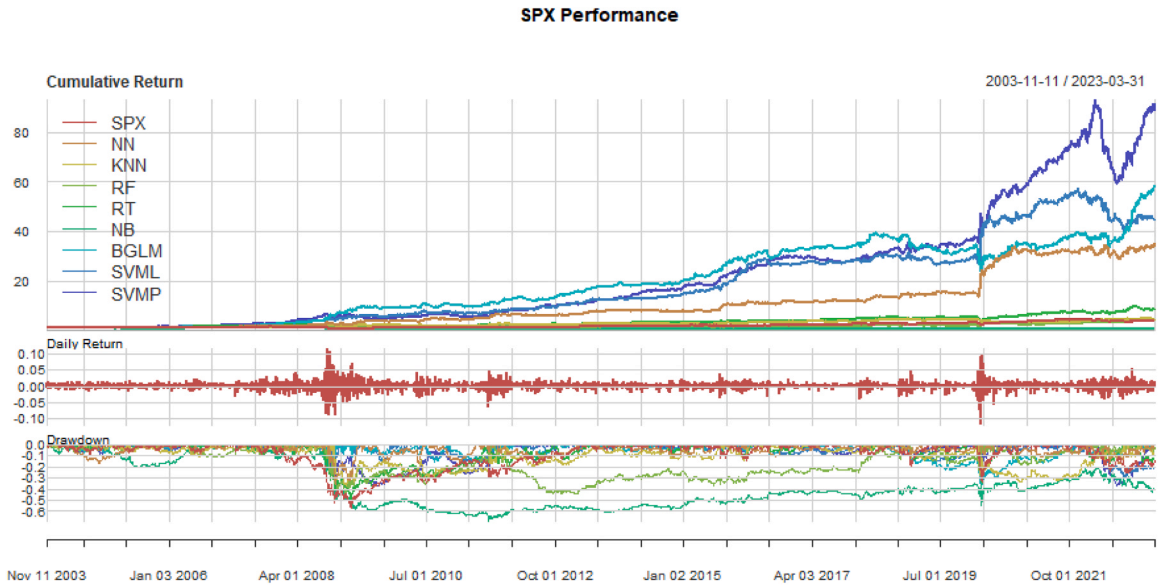


**Fig. 4.3.** Equity lines, daily returns and drawdown lines for S&P500 (USA). Note: Figure shows equity lines, daily returns and drawdown lines for every strategy constructed on S&P500 index in the period from 2003–11–11 to 2023–03–31.

matter.

Risk and return measures for S&P500 strategies are presented in Table 4.3. The best score of 1.05 for IR* measure was obtained by SVMP whereas the result for the benchmark strategy was 0.05. Polynomial Support Vector Machine model was therefore considered best performing from all analyzed strategies.

### 5.1.4. *Investment strategies comparison for CEE indices*

Results calculated for remaining CEE indices are described in this section in an aggregated manner. For the purpose of this research, strategies were compared only by the usage of the most decisive risk and return measure (IR*). Table 4.4 presents the strategies that received the highest IR* measure for each of the analyzed indices.

Naïve Bayes model was considered best performing for three out of six CEE indices with IR* value of 0.57 for SOFIX index (Bulgaria), 0.91 for OMXT index (Estonia) and 0.63 for OMXV index (Lithuania). SVML model which was dominant in case of DAX index was also considered best performing for OMXR index (Latvia) with the score of 0.44. RT model was best performing for BUX index (Hungary) with the IR* score of 0.16. Neural Networks achieved a highest IR* value of 0.10 for PX index (Czech Republic).

The next section summarizes the results from all analyzed indices and aims to propose a robust approach for the best performing model selection.

### 5.1.5. *Summary of investment strategies comparison for all indices*

As shown in the previous sections, results varied across all analyzed indices. There was no strategy that could be considered the best performing for all instruments. Table 4.5 similarly to Table 4.4 presents the strategies for which the highest IR* score was observed and the corresponding names of the models (strategies) for every index analyzed in this research. It is worth noting that the IR* values for benchmark buy-and-hold strategy were in all cases lower than those obtained by best performing machine learning models.

performed best for three less liquid indices i.e. SOFIX, OMXT and OMXV. Polynomial SVM was dominant for two more liquid

**Table 4.3**
Risk and return measures for S&P500 (USA).

| Measure | S&P500 | NN | KNN | RF | RT | NB | BGLM | SVML | SVMP |
|---|---|---|---|---|---|---|---|---|---|
| CAGR | 7.32% | 20.19% | 8.33% | 7.66% | 11.52% | -1.68% | 23.29% | 21.77% | 26.03% |
| Annual. Std Dev | 19.30% | 17.42% | 17.40% | 17.28% | 19.23% | 18.16% | 17.45% | 17.28% | 17.28% |
| Adj Sharpe | 0.3791 | 1.1590 | 0.4788 | 0.4435 | 0.5988 | -0.0926 | 1.3349 | 1.2598 | 1.5064 |
| MDD | 56.78% | 35.34% | 37.13% | 44.28% | 49.61% | 68.58% | 40.60% | 29.74% | 37.46% |
| IR* | 0.0489 | 0.6624 | 0.1074 | 0.0767 | 0.1390 | 0.0000 | 0.7658 | 0.9221 | 1.0470 |
| SoR | 0.0407 | 0.1131 | 0.0518 | 0.0483 | 0.0592 | 0.0000 | 0.1272 | 0.1205 | 0.1404 |
| CR | 0.1289 | 0.5715 | 0.2244 | 0.1731 | 0.2322 | 0.0000 | 0.5737 | 0.7319 | 0.6950 |

Note: Table shows risk and return measures for strategies constructed on S&P500 index. The first column represents buy-and-hold strategy. Presented measures include: CAGR, annualized standard deviation, adjusted Sharpe Ratio, Maximum Drawdown, IR* , Sortino Ratio and Calmar Ratio. Bolded font indicates the best performance measure for all tested methods.

**Table 4.4**
Best performing models and corresponding IR* measures for CEE indices.

| Index | BUX | PX | SOFIX | OMXR | OMXT | OMXV |
|---|---|---|---|---|---|---|
| Strategy | RT | NN | NB | SVML | NB | NB |
| IR* | 0.1587 | 0.1017 | 0.5690 | 0.4394 | 0.9127 | 0.6291 |

Note: Table shows IR* measure for the best performing strategies constructed on CEE indices: BUX, PX, SOFIX, OMXR, OMXT and OMXV.

**Table 4.5**
Best performing models and corresponding IR* measures for all analyzed indices.

| Index | WIG20 | DAX | SPX | BUX | PX | SOFIX | OMXR | OMXT | OMXV |
|---|---|---|---|---|---|---|---|---|---|
| Strategy | SVMP | SVML | SVMP | RT | NN | NB | SVML | NB | NB |
| IR* | 1.7234 | 0.3421 | 1.0470 | 0.1587 | 0.1017 | 0.5690 | 0.4394 | 0.9127 | 0.6291 |

Note: Table shows IR* measure for the best performing strategies constructed on all analyzed indices: WIG20, DAX, S&P500, BUX, PX, SOFIX, OMXR, OMXT and OMXV.

indices i.e. WIG20 and S&P500. Similar situation applies to Linear SVM which performed best for DAX and OMXR. Regression Tree model achieved the best results for BUX and Neural Networks model was considered the best for PX index.

In order to compare 9 analyzed strategies (8 machine learning models strategies plus the benchmark buy-and-hold strategy) across indices, a rank approach was introduced. Alternative comparison approaches were initially analyzed including different implementations of the ranking process (e.g. focusing on absolute differences between the values before the ranking process) but those were rejected due to the outcomes not providing any additional information and being more biased than in the proposed approach. For each index, strategies were ranked from 1 to 9 where 9 constitutes the highest score. For example, in case of WIG20, SVMP strategy had the highest IR* measure and received score equal to 9 in IR* category. Ranks were then averaged across all analyzed indices and presented in Table 4.6. The second column of the table corresponds to benchmark buy-and-hold strategy (B&H).

The best score of 6.28 for IR* averaged rank was obtained by SVML, the second best was observed for BGLM (6.11) and the third best for RT (5.78) whereas the result for the benchmark strategy was 5.00. Based on this analysis Linear Support Vector Machine (SVML) was considered the one producing the most robust results across all analyzed indices. The following sections describe sensitivity analysis performed to investigate if this conclusion changes when underlying models' parameters are altered.

### *5.2. Sensitivity to technical analysis indicators*

Technical analysis indicators served as inputs for machine learning models and therefore their levels are impacting trading signal generation in the analyzed strategies. Each technical indicator was calculated based on its underlying parameters which in all cases determine how many periods of observable stock index quotes were included in the calculation. Base parameters used in previous sections as well as parameters used for sensitivity analysis were described in the Technical Analysis Indicators section. Sensitivity analysis was performed in two scenarios. In the first scenario each base parameter was decreased by 1 which means that one period less was considered in computing technical indicators. In the second scenario each base parameter was increased by 1, which means that one period more was considered in computing the indicators. In order for the results to be comparable with the previous section, similar approach to presentation of the results was implemented.

#### *5.2.1. Scenario with decreased technical indicators' parameters*

Table 4.7 presents the strategies for which the highest IR* score was observed and corresponding names of the models (strategies) for each analyzed index. Naïve Bayes strategy once more performed the best for three indices i.e. SOFIX, OMXT and OMXV. Bayesian Generalized Linear Model strategy performed the best for two indices i.e. BUX and PX. SVML strategy had the best results for two indices i.e. S&P500 and OMXR. SVMP model was considered the best for WIG20 while RT was the best for DAX.

Ranks averaged across all indices analogous to those described for the base scenario are presented in Table 4.8. The best score of 6.67 for IR* averaged rank was obtained by BGLM, the second best was observed for SVML (6.44) and the third best for RT (5.44) whereas the result for the benchmark strategy was 5.22. Compared to the conclusion from base scenario, three best performing models remained the same with BGLM achieving best results in this sensitivity scenario.

**Table 4.6**
Ranked risk and return measures averaged across all analyzed indices.

| Measure | B&H | NN | KNN | RF | RT | NB | BGLM | SVML | SVMP |
|---|---|---|---|---|---|---|---|---|---|
| IR* | 5.00 | 4.89 | 4.00 | 4.89 | 5.78 | 3.78 | 6.11 | 6.28 | 4.28 |

Note: Table shows ranked risk and return measure IR* averaged across all analyzed indices. Bolded font indicates the best performance ranked measure for all tested methods.

**Table 4.7**
The best performing models and corresponding IR* measures for all analyzed indices in scenario with decreased technical indicators' parameters.

| Index | WIG20 | DAX | SPX | BUX | PX | SOFIX | OMXR | OMXT | OMXV |
|---|---|---|---|---|---|---|---|---|---|
| Strategy | SVMP | RT | SVML | BGLM | BGLM | NB | SVML | NB | NB |
| IR* | 1.3067 | 0.2357 | 0.6945 | 0.1403 | 0.1283 | 0.4461 | 0.4597 | 0.8492 | 1.1391 |

Note: Table shows IR* measure for the best performing strategies constructed on all analyzed indices in the sensitivity analysis scenario with decreased technical indicators' parameters.

**Table 4.8**
Ranked risk and return measures averaged across all analyzed indices in scenario with decreased technical indicators' parameters.

| Measure | B&H | NN | KNN | RF | RT | NB | BGLM | SVML | SVMP |
|---|---|---|---|---|---|---|---|---|---|
| IR* | 5.22 | 4.89 | 2.89 | 5.00 | 5.44 | 3.83 | 6.67 | 6.44 | 4.61 |

Note: Table shows ranked risk and return measure IR* averaged across all analyzed indices in the sensitivity analysis scenario with decreased technical indicators' parameters. Bolded font indicates the best performance ranked measure for all tested methods.

### 5.2.2. *Scenario with increased technical indicators' parameters*

Strategies for which the highest IR* score was observed in this scenario and the corresponding names of the models (strategies) are presented in Table 4.9. Naïve Bayes strategy again performed the best for three indices i.e. SOFIX, OMXT and OMXV. Linear Support Vector Machine (SVML) strategy achieved the best results also for three indices i.e. DAX, BUX and OMXR. Bayesian Generalized Linear Model strategy performed the best for S&P500. Polynomial version of SVM was considered the best for WIG20 index as in the base scenario. Neural Networks strategy was for the first time identified as the best performer for PX index.

Averaged ranks for increased parameters scenario are presented in Table 4.10. The best score of 6.61 for IR* averaged rank was obtained by BGLM, second best for SVML (6.28) strategies and the third best for RT (5.72) strategies whereas the result for the benchmark strategy was 5.11. Conclusion is the same as the conclusion from the scenario with decreased parameters – BGLM was the most robust strategy.

## 5.3. *Sensitivity to machine learning optimization metrics*

Every machine learning technique used in this research has a broad set of underlying hyperparameters which differs across the models. As there was no universal hyperparameter to be altered, the models were divided into three categories i.e. Neural Networks (considered a distinct category due to the implemented model algorithm not allowing for alteration of optimization metrics contrary to the other regression models), classification models (comprising of NB model) and regression models (comprising of the remaining six models). This sensitivity analysis exercise focused on assessing each model category separately in order to investigate if the strategies constructed from machine learning models' outputs are prone to changes in hyperparameters. Sensitivity assessment was based solely on IR* risk and return measure results as it was considered the most decisive in this research.

Model outputs as well as the corresponding trading signals were computed for all forms of the altered hyperparameter. It allowed for calculation of IR* measure in multiple scenarios for each of the analyzed stock indices. IR* was then compared and scenarios were ranked, with the highest score regarded as the best. Rank was then averaged across all analyzed indices.

### 5.3.1. *Sensitivity of Neural Networks*

In case of NN models, hyperparameter altered in sensitivity analysis was the activation function which was investigated in three different forms (scenarios) i.e. tansig (base function used in this research), sin and satlins as described in Machine Learning Techniques section, results of which are presented in Table 4.11.

Activation function sin received the best averaged rank (2.50) for all tested indices, satlins received 1.89 and tansig received 1.61 score. Those results can be interpreted in the following manner: by altering the activation function to sin, on average the NN models produce returns with higher IR* measure than those obtained from the employment of satlins and tansig functions.

### 5.3.2. *Sensitivity of classification models*

Naïve Bayes model is the only classification model described in this research. Hyperparameter that was altered in sensitivity analysis was the optimization metric investigated in two different versions (scenarios) i.e. accuracy (base metric used in this research)

**Table 4.9**
The best performing models and corresponding IR* measures for all analyzed indices in scenario with increased technical indicators' parameters.

| Index | WIG20 | DAX | SPX | BUX | PX | SOFIX | OMXR | OMXT | OMXV |
|---|---|---|---|---|---|---|---|---|---|
| Strategy | SVMP | SVML | BGLM | SVML | NN | NB | SVML | NB | NB |
| IR* | 1.6824 | 0.3437 | 0.9686 | 0.2854 | 0.0314 | 0.3544 | 0.5180 | 0.7199 | 0.5478 |

Note: Table shows IR* measure for the best performing strategies constructed on all analyzed indices in the sensitivity analysis scenario with increased technical indicators' parameters.

**Table 4.10**
Ranked risk and return measures averaged across all analyzed indices in scenario with increased technical indicators' parameters.

| Measure | B&H | NN | KNN | RF | RT | NB | BGLM | SVML | SVMP |
|---|---|---|---|---|---|---|---|---|---|
| IR* | 5.11 | 5.06 | 2.78 | 5.22 | 5.72 | 4.06 | 6.61 | 6.28 | 4.17 |

Note: Table shows ranked risk and return measure IR* averaged across all analyzed indices in the sensitivity analysis scenario with increased technical indicators' parameters. Bolded font indicates the best performance ranked measure for all tested methods.

**Table 4.11**
Ranked IR* measure averaged across all analyzed indices for Neural Networks in three activation function scenarios.

| Activation Function | Neural Networks |
|---|---|
| Tansig | 1.61 |
| Sin | 2.50 |
| Satlins | 1.89 |

Note: Table shows ranked IR* measure for tansig, sin and satlins activation functions applied in NN models averaged across all analyzed indices. Bolded font indicates the best performance ranked measure for all tested methods.

and kappa as described in Machine Learning Techniques section, results of which are presented in Table 4.12.

Kappa metric achieved a higher score of 1.56 compared to 1.44 obtained by accuracy metric. This means that Naïve Bayes model on average produces higher IR* when kappa metric is applied.

### *5.3.3. Sensitivity of regression models*

Regression models category comprises of six models i.e. KNN, RF, RT, BGLM, SVML and SVMP. As in the case of classification models the hyperparameter altered in sensitivity analysis was the optimization metric. As described in Machine Learning Techniques section, optimization metrics for regression models differed from those that could be applied for Neural Networks and classification models. Metrics were therefore investigated in three different versions (scenarios) i.e. RMSE (root mean square error – base metric used in this research), Rsquared (coefficient of determination) and MAE (mean absolute error), results of which are presented in Table 4.13.

computed with the RMSE metric generated on average the best IR* values (2.33 score). Polynomial Support Vector Machine and Random Forest models obtained the highest averaged rank of 2.33 and 2.22 respectively when computed with the Rsquared metric which means that on average SVMP and RF models produce higher IR* when Rsquared metric is applied. For Regression Tree model, computation with the MAE metric generated on average the best IR* values (2.33 score). In case of Bayesian Generalized Linear Model and Linear Support Vector Machine, models obtained the same averaged rank (2.00) for every metric which means that they are insensitive to optimization metric alteration.

## *5.4. Sensitivity to the length of in-sample and out-of-sample windows*

The length of in-sample and out-of-sample windows impacts the construction and application of each model, and it is commonly selected using arbitrary methods which is why it is important to investigate the sensitivity of these decisions. For the purpose of this sensitivity analysis, lengths of in-sample and out-of-sample windows were decreased and increased which constitutes four additional distinct sensitivity scenarios.

The length of in-sample determines how much data is used to build a model. Base scenario assumes in-sample to incorporate 200 trading days. Adding more days (observations) to in-sample period in general terms means adding more, less recent observations to the sample. Shortening in-sample period means deleting less recent observations which in theory has a potential to improve the results but at the same time it can make the model less robust or more volatile which is a downside for model performance over long time periods. For the purpose of sensitivity analysis with regard to in-sample period, initial size of 200 observations was multiplied/divided by 2

**Table 4.12**
Ranked IR* measure averaged across all analyzed indices for classification models in two optimization metric scenarios.

| Optimization metric | Naïve Bayes |
|---|---|
| Accuracy | 1.44 |
| Kappa | 1.56 |

Note: Table shows ranked IR* measure for accuracy and kappa optimization metrics applied in NB models averaged across all analyzed indices. Bolded font indicates the best performance ranked measure for all tested methods.

**Table 4.13**
Ranked IR* measure averaged across all analyzed indices for regression models in three optimization metric scenarios.

| Optimization metric | KNN | RF | RT | BGLM | SVML | SVMP |
|---|---|---|---|---|---|---|
| RMSE | 2.33 | 1.67 | 1.89 | 2.00 | 2.00 | 1.83 |
| Rsquared | 2.00 | 2.22 | 1.78 | 2.00 | 2.00 | 2.33 |
| MAE | 1.67 | 2.11 | 2.33 | 2.00 | 2.00 | 1.83 |

Note: Table shows ranked IR* measure for RMSE, Rsquared and MAE optimization metrics applied in regression models averaged across all analyzed indices. Bolded font indicates the best performance ranked measure for all tested methods.

which produced scenario with in-sample period having 400 and 100 observations respectively.

For out-of-sample, the base scenario assumed 20 trading days. This period is used to calculate forecasts using the model built on in-sample period. Adding more days to out-of-sample means adding observations that are further away (and can deviate more) from in-sample period. Shortening out-of-sample means forecasting on a smaller number of observations and re-estimating the model more frequently to capture more recent moves observed in the market which in theory can make the predictions more precise (up to date). For the purpose of sensitivity analysis with regard to out-of-sample period, initial size of 20 observations was multiplied/divided by 2 which produced scenario with out-of-sample period having 40 and 10 observations respectively. Four different sensitivity scenarios are described in subsections below. For the results to be comparable with the previous sections, a similar approach to presentation of the results was implemented.

#### *5.4.1. Scenario with decreased length of in-sample*

Table 4.14 presents the strategies for which the highest IR* score was observed and corresponding names of the models (strategies) for each analyzed index. Random Forest strategy performed the best for four indices i.e. DAX, BUX, PX and OMXV. For each of the remaining five indices, there was a different top performing strategy resulting from this scenario. Polynomial Support Vector Machine model was considered the best for WIG20, Linear Support Vector Machine for S&P500, Naïve Bayes for SOFIX, Neural Networks for OMXR and Regression Tree for OMXT.

Ranks averaged across all indices analogous to those described for the base scenario are presented in Table 4.15. The best score of 6.94 for IR* averaged rank was obtained by SVML, the second best was observed for RF (6.78) and the third best for BGLM (6.44) whereas the result for the benchmark strategy was 3.89. As in the case of base scenario, Linear Support Vector Machine received the most robust results across all analyzed indices.

#### *5.4.2. Scenario with increased length of in-sample*

Strategies for which the highest IR* score was observed in this scenario and the corresponding names of the models (strategies) are presented in Table 4.16. Polynomial Support Vector Machine (SVML) strategy achieved the best results for three most liquid indices i.e. WIG20, DAX and S&P500. Naïve Bayes strategy again performed the best for three indices i.e. SOFIX, OMXT and OMXV. Regression Tree strategy performed the best for BUX and PX. Linear version of SVM was considered the best for OMXR index.

Averaged ranks for this scenario are presented in Table 4.17. For the first time across ranked results, first place is divided between two strategies, one of which being the benchmark buy-and-hold and second one being Linear Support Vector Machine with scores of 5.44. The third best score of 5.39 for IR* averaged rank was obtained by Random Forest model. Adding more days to the in-sample window which translates into more, less recent information taken into account in construction of the models, produces results that are in most cases worse than the benchmark in this scenario. Nevertheless, SVML model once again was considered as being (one of) the best models which is consistent with base scenario results.

#### *5.4.3. Scenario with decreased length of out-of-sample*

Table 4.18 presents the strategies for which the highest IR* score was observed and corresponding names of the models (strategies) for each analyzed index. Naïve Bayes strategy once more performed the best for three indices i.e. SOFIX, OMXT and OMXV. Bayesian Generalized Linear Model strategy performed best for two indices i.e. S&P500 and OMXR. Polynomial Support Vector Machine model was considered the best for WIG20, Linear Support Vector Machine for DAX, Random Forest for BUX and Neural Networks for PX.

Ranks averaged across all indices analogous to those described for the base scenario are presented in Table 4.19. The best score of 6.11 for IR* averaged rank was obtained by BGLM, the second best was observed for RT (5.83) and the third best for SVML (5.67) whereas the result for the benchmark strategy was 4.44. Compared to the conclusion from base scenario, three top performing models remained the same with BGLM achieving best results in this sensitivity scenario.

**Table 4.14**
The best performing models and corresponding IR* measures for all analyzed indices in scenario with decreased length of in-sample.

| Index | WIG20 | DAX | SPX | BUX | PX | SOFIX | OMXR | OMXT | OMXV |
|---|---|---|---|---|---|---|---|---|---|
| Strategy | SVMP | RF | SVML | RF | RF | NB | NN | RT | RF |
| IR* | 1.4835 | 0.3001 | 0.5877 | 0.2060 | 0.1130 | 0.0823 | 1.3063 | 0.1780 | 0.4205 |

Note: Table shows IR* measure for the best performing strategies constructed on all analyzed indices in the sensitivity analysis scenario with decreased length of in-sample.

**Table 4.15**
Ranked risk and return measures averaged across all analyzed indices in scenario with decreased length of in-sample.

| Measure | B&H | NN | KNN | RF | RT | NB | BGLM | SVML | SVMP |
|---|---|---|---|---|---|---|---|---|---|
| IR* | 3.89 | 5.33 | 2.94 | 6.78 | 4.67 | 3.50 | 6.44 | 6.94 | 4.50 |

Note: Table shows ranked risk and return measure IR* averaged across all analyzed indices in the sensitivity analysis scenario with decreased length of in-sample. Bolded font indicates the best performance ranked measure for all tested methods.

**Table 4.16**
The best performing models and corresponding IR* measures for all analyzed indices in scenario with increased length of in-sample.

| Index | WIG20 | DAX | SPX | BUX | PX | SOFIX | OMXR | OMXT | OMXV |
|---|---|---|---|---|---|---|---|---|---|
| Strategy | SVMP | SVMP | SVMP | RT | RT | NB | SVML | NB | NB |
| IR* | 2.2027 | 0.1093 | 2.0677 | 0.2335 | 0.0370 | 1.1702 | 0.3346 | 1.2902 | 1.6374 |

Note: Table shows IR* measure for the best performing strategies constructed on all analyzed indices in the sensitivity analysis scenario with increased length of in-sample.

**Table 4.17**
Ranked risk and return measures averaged across all analyzed indices in scenario with increased length of in-sample.

| Measure | B&H | NN | KNN | RF | RT | NB | BGLM | SVML | SVMP |
|---|---|---|---|---|---|---|---|---|---|
| IR* | 5.44 | 5.06 | 3.39 | 5.39 | 5.17 | 4.72 | 5.28 | 5.44 | 5.11 |

Note: Table shows ranked risk and return measure IR* averaged across all analyzed indices in the sensitivity analysis scenario with increased length of in-sample. Bolded font indicates the best performance ranked measure for all tested methods.

**Table 4.18**
The best performing models and corresponding IR* measures for all analyzed indices in scenario with decreased length of out-of-sample.

| Index | WIG20 | DAX | SPX | BUX | PX | SOFIX | OMXR | OMXT | OMXV |
|---|---|---|---|---|---|---|---|---|---|
| Strategy | SVMP | SVML | BGLM | RF | NN | NB | BGLM | NB | NB |
| IR* | 3.2772 | 0.5230 | 2.4801 | 0.2347 | 0.1755 | 0.3976 | 1.5034 | 0.2408 | 0.7222 |

Note: Table shows IR* measure for the best performing strategies constructed on all analyzed indices in the sensitivity analysis scenario with decreased length of out-of-sample.

**Table 4.19**
Ranked risk and return measures averaged across all analyzed indices in scenario with decreased length of out-of-sample.

| Measure | B&H | NN | KNN | RF | RT | NB | BGLM | SVML | SVMP |
|---|---|---|---|---|---|---|---|---|---|
| IR* | 4.44 | 5.33 | 4.11 | 5.06 | 5.83 | 3.72 | 6.11 | 5.67 | 4.72 |

Note: Table shows ranked risk and return measure IR* averaged across all analyzed indices in the sensitivity analysis scenario with decreased length of out-of-sample. Bolded font indicates the best performance ranked measure for all tested methods.

#### *5.4.4. Scenario with increased length of out-of-sample*

Strategies for which the highest IR* score was observed in this scenario and the corresponding names of the models (strategies) are presented in Table 4.20. Naïve Bayes strategy again performed the best for three indices i.e. SOFIX, OMXT and OMXV. Linear Support Vector Machine strategy achieved the best results for two indices i.e. S&P500 and OMXR. Polynomial Linear Support Vector Machine was considered the best for WIG20, Bayesian Generalized Linear Model for DAX, Random Forest for BUX and Neural Networks for PX.

Averaged ranks for this scenario are presented in Table 4.21. The first place is divided between two models i.e. SVML and BGLM with the score of 6.78, the third best result was observed for RF (5.44) strategy whereas the result for the benchmark strategy was 4.78. Once again, SVML and BGLM models were among the best performing strategies in this scenario.

**Table 4.20**
The best performing models and corresponding IR* measures for all analyzed indices in scenario with increased length of out-of-sample.

| Index | WIG20 | DAX | SPX | BUX | PX | SOFIX | OMXR | OMXT | OMXV |
|---|---|---|---|---|---|---|---|---|---|
| Strategy | SVMP | BGLM | SVML | RF | NN | NB | SVML | NB | NB |
| IR* | 0.4234 | 0.3609 | 0.4565 | 0.1961 | 0.0184 | 0.6532 | 0.4483 | 1.8011 | 0.3703 |

Note: Table shows IR* measure for the best performing strategies constructed on all analyzed indices in the sensitivity analysis scenario with increased length of out-of-sample.

### 5.5. Sensitivity to different market conditions

Market conditions are constantly changing and among other factors, it depends on the selected frequency if we observe rising or declining price trends for particular instrument. Nevertheless, global trends which impact the biggest economies and most liquid market indices can be defined and divided into two main types i.e. bull and bear markets. Bull markets are periods of raising market prices while bear markets constitute of periods with (more rapidly) declining prices.

As it is impossible to know a priori in which market state we currently are when building a model, this sensitivity analysis is an a posteriori investigation of how the base scenario model behaves in different market conditions which means that in-sample windows and construction of the models are the same as in the base scenario. The difference here is how the forecasts i.e. out-of-sample periods were utilized. For the scenario in which bull (bear) market conditions are analyzed, all the daily forecasts calculated in periods considered to be a part of bear (bull) market are neutralized i.e. signal is overridden to 0 which means to hold an instrument.

The last objective in this exercise was to define bull and bear periods. Bull periods constituted of the following date ranges: from 01.01.2002 (beginning of analyzed data) to 08.10.2007, from 10.03.2009 to 18.02.2020 and from 24.03.2020 to 02.01.2022. Bear market was defined in the remaining periods: from 09.10.2007 to 09.03.2009 (great financial crisis of 2007–2009), from 19.02.2020 to 23.03.2020 (COVID-19 pandemic crisis) and from 03.01.2022 to 31.03.2023 (drawdowns that started in 2021 which reflected market concerns related to the growing inflation and the accompanying increases in interest rates till the last day of analyzed time period).

#### 5.5.1. Scenario with bull market conditions

Table 4.22 presents the strategies for which the highest IR* score was observed and corresponding names of the models (strategies) for each analyzed index. For the first time among all the sensitivity analyses, buy-and-hold (B&H) strategy was considered the best and it was observed for five indices i.e. BUX, PX, SOFIX, OMXT and OMXV. Polynomial Support Vector Machine model was considered the best for WIG20, Regression Tree for DAX, Neural Networks for S&P500 and Linear Support Vector Machine for OMXR.

Ranks averaged across all indices analogous to those described for the base scenario are presented in Table 4.23. The best score of 7.78 for IR* averaged rank was obtained by B&H, the second best was observed for RT (6.67) and the third best for SVML (6.17). Buy-and-hold strategy achieved the best score which means that on average, there was no Machine Learning model that could outperform the benchmark in bull market conditions.

#### 5.5.2. Scenario with bear market conditions

Strategies for which the highest IR* score was observed in this scenario and the corresponding names of the models (strategies) are presented in Table 4.24. Naïve Bayes strategy again performed the best for three indices i.e. SOFIX, OMXT and OMXV. Bayesian Generalized Linear Model strategy achieved the best results for DAX and S&P500. Neural Networks strategy performed the best for WIG20 and PX. Random Forest was considered the best for BUX and K Nearest Neighbor for OMXR.

Averaged ranks for this scenario are presented in Table 4.25. For the first time in all sensitivity scenarios, Neural Networks achieved the first place with a score of 6.72. The second best score of 6.11 was observed for BGLM and the third best score of 5.94 for RF. While Bayesian Generalized Linear Model achieved a high score, results for Neural Networks strategies suggest that on average, this model is the most appropriate in bear market conditions.

### 5.6. Summary of the empirical research

Researching a set of nine different investment strategies allowed to obtain answers to the hypotheses stated in the research. With the employment of risk and return measures such as adjusted Information Ratio as a selection criteria, it was possible to select the best performing strategy.

Models were fitted and applied on dynamic windows of in-sample and out-of-sample subsets with the dates ranging from 2002 to 2023 for multiple global and CEE stock indices. For WIG20 index (Poland), the Polynomial Support Vector Machine strategy achieved IR* of 1.72 compared to 0.0001 achieved by the benchmark. For DAX, the best performing strategy was based on Linear Support Vector Machine with 0.34 IR* while benchmark achieved 0.06 IR* . In case of S&P500, the Polynomial Support Vector Machine strategy obtained the best IR* amounting to 1.05 while benchmark buy-and-hold strategy resulted in 0.05 IR* .

As part of the research, results for six CEE stock indices were also analyzed. In all cases, a strategy superior to the benchmark was identified. Naïve Bayes strategy was considered as the best performing for three indices with IR* value of 0.57 for SOFIX index (Bulgaria), 0.91 for OMXT index (Estonia) and 0.63 for OMXV index (Lithuania). SVML strategy obtained the highest IR* value (0.44) in case of OMXR index (Latvia). RT strategy performed the best for BUX index (Hungary) with IR* of 0.16. NN strategy was considered the best performing for PX index (Czech Republic) with IR* of 0.10.

Strategies were ranked according to the value of IR* for each index separately and the ranks were then averaged among indices. The best score of 6.28 for IR* averaged rank was obtained by SVML, the second best was observed for BGLM (6.11) and the third best for RF (5.78) whereas the result for the benchmark strategy was 5.00.

First of all, sensitivity analysis was conducted in order to investigate whether a change (decrease and increase) in the parameters underlying technical analysis indicators which serve as inputs to the models would impact the results. Three best performing models remained the same as in the base scenario with BGLM as the best model instead of SVML. Based on this analysis, Bayesian Generalized Linear Model (BGLM) was considered to be producing the most robust results across all analyzed indices in these sensitivity analysis scenarios.

The second part of sensitivity analysis investigated the impact of changing the hyperparameters underlying machine learning

**Table 4.21**
Ranked risk and return measures averaged across all analyzed indices in scenario with increased length of out-of-sample.

| Measure | B&H | NN | KNN | RF | RT | NB | BGLM | SVML | SVMP |
|---|---|---|---|---|---|---|---|---|---|
| IR* | 4.78 | 5.11 | 3.11 | 5.44 | 4.06 | 4.33 | 6.78 | 6.78 | 4.61 |

Note: Table shows ranked risk and return measure IR* averaged across all analyzed indices in the sensitivity analysis scenario with increased length of out-of-sample. Bolded font indicates the best performance ranked measure for all tested methods.

**Table 4.22**
The best performing models and corresponding IR* measures for all analyzed indices in scenario with bull market conditions.

| Index | WIG20 | DAX | SPX | BUX | PX | SOFIX | OMXR | OMXT | OMXV |
|---|---|---|---|---|---|---|---|---|---|
| Strategy | SVMP | RT | NN | B&H | B&H | B&H | SVML | B&H | B&H |
| IR* | 2.3086 | 0.4408 | 1.0665 | 0.3865 | 0.2209 | 0.3505 | 0.9953 | 0.8464 | 0.8810 |

Note: Table shows IR* measure for the best performing strategies constructed on all analyzed indices in the sensitivity analysis scenario with bull market conditions.

**Table 4.23**
Ranked risk and return measures averaged across all analyzed indices in scenario with bull market conditions.

| Measure | B&H | NN | KNN | RF | RT | NB | BGLM | SVML | SVMP |
|---|---|---|---|---|---|---|---|---|---|
| IR* | 7.78 | 4.56 | 3.44 | 3.56 | 6.67 | 3.44 | 4.78 | 6.17 | 4.61 |

Note: Table shows ranked risk and return measure IR* averaged across all analyzed indices in the sensitivity analysis scenario with bull market conditions. Bolded font indicates the best performance ranked measure for all tested methods.

**Table 4.24**
The best performing models and corresponding IR* measures for all analyzed indices in scenario with bear market conditions.

| Index | WIG20 | DAX | SPX | BUX | PX | SOFIX | OMXR | OMXT | OMXV |
|---|---|---|---|---|---|---|---|---|---|
| Strategy | NN | BGLM | BGLM | RF | NN | NB | KNN | NB | NB |
| IR* | 0.0567 | 0.0205 | 0.1040 | 0.3560 | 0.0584 | 0.3391 | 0.0019 | 0.2119 | 0.1606 |

Note: Table shows IR* measure for the best performing strategies constructed on all analyzed indices in the sensitivity analysis scenario with bear market conditions.

**Table 4.25**
Ranked risk and return measures averaged across all analyzed indices in scenario with bear market conditions.

| Measure | B&H | NN | KNN | RF | RT | NB | BGLM | SVML | SVMP |
|---|---|---|---|---|---|---|---|---|---|
| IR* | 3.28 | 6.72 | 5.06 | 5.94 | 3.28 | 5.28 | 6.11 | 5.00 | 4.33 |

Note: Table shows ranked risk and return measure IR* averaged across all analyzed indices in the sensitivity analysis scenario with bear market conditions. Bolded font indicates the best performance ranked measure for all tested methods.

models. As there was no universal hyperparameter to be altered, activation function was chosen for Neural Networks strategies and optimization metric was chosen for the remaining strategies. Linear Support Vector Machine and Bayesian Generalized Linear Model were insensitive to optimization metrics alteration.

For the third part of sensitivity analysis, changes to sampling approach were investigated. In-sample and out-of-sample period was decreased and increased in length which constituted four distinct scenarios. Although the results differed between scenarios, Linear Support Vector Machine and Bayesian Generalized Linear Model were occurring as top performing models most frequently with SVML as the one considered the best and most robust on average.

The fourth and the last part of sensitivity analysis considered how the models constructed in base scenario would behave in bull and bear market conditions. For the bull market conditions, buy-and-hold strategy was the dominant one which means that there was no single Machine Learning model that could outperform the benchmark. For the bear market conditions, while Bayesian Generalized Linear Model achieved a high score, the Neural Networks model produced on average the best results across all analyzed indices.

## 6. Conclusions

New technologies allowed for transfer of stock exchange trading from the trading floor to the computer screen, and also opened the possibility of automating the investment process i.e. execution of transactions with limited or even without human intervention. Automated investment strategies are now widely used by hedge funds and rely among others on rules derived from technical analysis.

Technical analysis is a rich set of tools supporting investment decision making. An investor can use them directly in the construction of buy and sell signals as well as indirectly by treating them as inputs to more sophisticated models. Employment of machine learning techniques allows to generate weights for every technical analysis indicator used and therefore producing trading signals based on both the levels and the weights estimated for the technical indicators.

A set of five technical analysis indicators was analyzed in this research: Simple Moving Average (SMA), Moving Average Convergence Divergence (MACD), Stochastic Oscillator (STOCH), Relative Strength Index (RSI) and Williams' Percent Range (WPR) as proposed by Dash and Dash (2016). Technical indicators were then used as an input to eight machine learning models. The techniques analyzed were: Neural Networks (NN), K Nearest Neighbor (KNN), Regression Tree (RT), Random Forest (RF), Naïve Bayes (NB), Bayesian Generalized Linear Model (BGLM), Linear Support Vector Machine (SVML) and Polynomial Support Vector Machine (SVMP).

The purpose of this study was to investigate the profitability of machine learning-based quantitative investment strategies. Performance of the strategies was determined by comparison of the risk and return measures among models and the benchmark buy-and-hold returns. annualized rate of return (CAGR), annualized standard deviation of returns, adjusted Sharpe Ratio (SR), maximum drawdown (MDD) and the adjusted Information Ratio (IR*) were used as the aforementioned risk and return measures with the IR* considered as the most significant. Sortino Ratio (SoR) and Calmar Ratio (CR) were additionally presented for informational purposes.

The research was based on the current scientific achievements describing the mechanisms of generating trading signals from machine learning models employing technical indicators as inputs. As part of the extension of this branch of science, this research was conducted on the Polish stock market index WIG20, two highly liquid equity indices: DAX (Germany) and S&P500 (USA) as well as the indices of six Central and Eastern European countries: BUX (Hungary), PX (Czech Republic), SOFIX (Bulgaria), OMXR (Latvia), OMXT (Estonia) and OMXV (Lithuania). Data used for the calculations included the daily High, Low and Close prices of the indices in 2002–2023 period thus including the 2007–2009 great financial crisis, COVID-19 pandemic crisis and drawdowns that started at the end of 2021 which reflected market concerns related to the growing inflation and the accompanying increases in interest rates. Models were fitted in dynamic in-sample window and applied to out-of-sample subsets separately for each of the analyzed indices.

Results showed that in the case of each index, machine learning techniques-based strategies achieved better returns than the benchmarks. The best performing (based on IR* measure) strategy was constructed from Polynomial Support Vector Machine model in case of WIG20 index (Poland), Linear Support Vector Machine model for DAX (Germany) and Polynomial Support Vector Machine for S&P500 (USA). Comparison of those results with the ones obtained for six CEE indices showed that on average the Linear Support Vector Machine strategy generated the best risk adjusted returns.

In order to investigate the robustness of analyzed quantitative strategies, the sensitivity analysis was conducted. First of all, investigation was conducted whether the change (decrease and increase) in the parameters underlying technical analysis indicators which served as inputs to the models would impact the results. Altering the parameters in both directions showed that on average the Bayesian Generalized Linear Model (BGLM) produced the most robust results across analyzed indices. The second part of sensitivity analysis investigated the impact of changing the hyperparameters underlying the machine learning models. In case of Linear Support Vector Machine and Bayesian Generalized Linear Model, results showed that returns generated by these models are insensitive to optimization metric alteration. The third sensitivity analysis challenged the length of in-sample and out-of-sample windows and indicated that Linear Support Vector Machine and Bayesian Generalized Linear Model were the most robust strategies. The fourth and last sensitivity analysis investigated which models are the best performing depending on market conditions. For bull market, there was no ML model that could outperform the benchmark buy-and-hold strategy while for the bear market, Neural Networks and Bayesian Generalized Linear Model achieved the best results.

The returns obtained from the trading signals generated by machine learning models indicated that quantitative investment strategies achieved better performance measured by adjusted Information Ratio than the benchmark buy-and-hold strategies for all of the analyzed stock market indices, and therefore the main research hypothesis (RH1: *Active quantitative investment strategies based on the signals generated by machine learning models result in higher risk-adjusted returns than the buy-and-hold benchmark strategy*) cannot be rejected. It should be stated that the range of financial instruments available for investment is very wide, and in order to unequivocally assess this hypothesis, a larger number of instruments should be tested in the future, especially focusing on various asset classes rather than only the equity type. The additional research hypotheses stated in this research were formulated with the aim of assessing the best-performing machine learning techniques across the analyzed indices. The analysis of each index allowed to select the best-performing strategy, which differed among indices, but at the same time, it was observed that on average the Linear Support Vector Machine produced the best risk-adjusted returns, thus the second and the third hypotheses were rejected (RH2: *Neural Networks generate the best (with regard to risk-adjusted returns) investment signals compared to other machine learning techniques used in the research*; RH3: *The very same machine learning strategy is considered best performing for all analyzed stock market indices*). The sensitivity analysis results showed that the fourth hypothesis also needs to be rejected (RH4: *Returns obtained from signals generated by machine learning techniques are resistant to: changes in hyperparameters underlying the models, changes in parameters underlying the technical analysis indicators, changes in sampling and differing market conditions*). On a strategy level, results changed in each analyzed sensitivity scenario for most of the analyzed models. On average, however, the Bayesian Generalized Linear Model generated the best results in sensitivity analysis scenarios in which the technical analysis indicators' parameters were altered. In the case of altering the machine learning models' hyperparameters, the SVML and BGLM models were insensitive to changes. Sensitivity to changes in sampling approach indicated that SVML and BGLM models were the most robust ones. In the case of investigating different market conditions, buy-and-hold strategy was the best for bull markets while NN and BGLM models were the best in bear markets.

In this place, it is worth noting that these findings go beyond just the technical results of an academic publication and can constitute some guidance to market practitioners (quant researchers, analysts, and decision makers) working in the financial industry (hedge funds, mutual funds, pension funds, or insurance companies) in the process of investment strategies development. The results of this study, summarized in Table 4.6 for all equity indices and all tested models, indicate which models (SVML, BGLM and RT) should be taken into account as preliminary results and suggestions for future enhancement while designing and testing such investment strategies on equity markets. At the same time, the architecture of testing (walk-forward approach with numerous in-sample and out -of -sample windows) and quite thorough sensitivity analysis (feature engineering in the input layer, type of loss function, and activation function, or sensitivity to the type of market trend) may provide some direction for government regulators with regards to the industry standard of testing of investment models before their implementation in real asset management processes.

In the end, it is worth noting that we are aware of the fact that the study based on daily data for equity indices limits its conclusions to the selected frequency, chosen time period, and type of assets. Therefore, in order to obtain a comprehensive overview of the topics presented in this research, future research could be expanded by including financial instruments from other asset classes, e.g. stocks, bonds, options, or currency futures. An interesting research direction would also be to investigate how the results would change if we focused on other types of ML models or some techniques based on NLP.

## Ethics approval

Not applicable.

## Consent to participate

Not applicable.

## Availability of data and material

Not applicable.

## Note

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Code availability

Not applicable.

## Funding

This research received no external funding.

## CRediT authorship contribution statement

**Jan Grudniewicz:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft preparation, Writing – review & editing, Visualization, Project administration. **Robert Ślepaczuk:** Conceptualization, Methodology, Validation, Investigation, Writing – original draft preparation, Writing – review & editing, Supervision, Project administration.

## Declaration of Competing Interest

The authors declare no conflict of interest

## Data Availability

Data will be made available on request.

## References

Akyildirim, E., Cepni, O., Corbet, S., Uddin, G.S., 2021. Forecasting mid-price movement of Bitcoin futures using machine learning. Ann. Oper. Res. 1–32.
Altman, N.S., 1991. An Introduction to Kernel and Nearest Neighbor Nonparametric Regression. Cornell University.
Anghel, D.G., 2021. A reality check on trading rule performance in the cryptocurrency market: Machine learning vs. technical analysis. Financ. Res. Lett. 39, 101655.
Anson, M., Fabozzi, F., Jones, F., 2010. The Handbook of Traditional and Alternative Investment Vehicles: Investment Characteristics and Strategies. John Wiley & Sons,.
Appel, G., 2005. Technical Analysis Power Tools for Active Investors. FT Press,, New Jersey.
Bacon, C., 2021. Practical Risk-adjusted Performance Measurement. John Wiley & Sons.
Będowska-Sójka, B., Kliber, A., 2022. Can cryptocurrencies hedge oil price fluctuations? a pandemic perspective. Energy Econ. 115, 106360.
Beechey M., Gruen, D., Vickery J., 2000. The Efficiency Market Hypothesis: A Survey. Sydney: Reserve Bank of Australia.

Bonato, M., Çepni, O., Gupta, R., Pierdzioch, C., 2022. El Niño, La Niña, and forecastability of the realized variance of agricultural commodity prices: Evidence from a machine learning approach. J. Forecast. 1–17.
Boser B.E., Guyon I.M., Vapnik V.N., 1992. A training algorithm for optimal margin classifiers. In Proceedings of the fifth annual workshop on Computational learnin‘ theory, 144–152.
Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32.
Breiman L., Friedman J.H., Olshen R.A., Stone C.J., 1984. Classification and Regression Trees. Wadsworth & Brooks. Cole Statistics/Probability Series.
Brock, W., Lakonishok, J., LeBaron, B., 1922. Simple technical trading rules and the stochastic properties of stock returns. J. Financ. 47 (5), 1731–1764.
Cepni, O., Gupta, R., Onay, Y., 2022a. The role of investor sentiment in forecasting housing returns in China: a machine learning approach. J. Forecast. 41 (8), 1725–1740.
Çepni, O., Gupta, R., Pienaar, D., Pierdzioch, C., 2022b. Forecasting the realized variance of oil-price returns using machine learning: is there a role for US state-level uncertainty? Energy Econ. 114, 106229.
Chen, W.-H., Shih, J.-Y., Wu, S., 2006. Comparison of support-vector machines and back propagation neural networks in forecasting the six major Asian stock markets. Int. J. Electron. Financ. 1 (1), 49–67.
Colianni, S., Rosales, S., Signorotti, M., 2015. Algorithmic trading of cryptocurrency based on twitter sentiment analysis. CS229 Proj. 1 (5), 1–4.
Cortes, C., Vapnik, V.N., 1995. Support-vector networks. Mach. Learn. 20, 273–297.
Dash, R., Dash, P.K., 2016. A hybrid stock trading framework integrating technical analysis with machine learning techniques. J. Financ. Data Sci. 2 (1), 42–57.
Dempster A.P., Laird N.M., Rubin D.B. Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society: series B, 39(1), 1–38.
Fama, E., 1970. Efficient Capital Markets: a review of theory and empirical work. The. J. Financ. 25, 383–417.
Feller, W., 1991. An Introduction to Probability Theory and its Applications, Volume 2. John Wiley & Sons.
Fiszeder, P., Orzeszko, W., 2021. Covariance matrix forecasting using support vector regression. Appl. Intell. 51 (10), 7029–7042.
Gerlein, E., McGinnity, M., Belatreche, A., Coleman, S., 2016. Evaluating machine learning classification for financial trading: an empirical approach. Expert Syst. Appl. 54, 193–207.
González-Sánchez, M., 2022. Factorial asset pricing models using statistical anomalies. Res. Int. Bus. Financ. 60, 101595.
Goutte, S., Le, H.V., Liu, F., von Mettenheim, H.J., 2023. Deep learning and technical analysis in cryptocurrency market. Financ. Res. Lett., 103809
Han, J., Kamber, M., Pei, J., 2011. Data mining: concepts and techniques. Concepts Tech. 3.
Huang, G., Zhu, Q., Siew, C., 2006. Extreme learning machine: theory and applications. Neurocomputing 70 (1–3), 489–501.
Huang, W., Nakamori, Y., Wang, S., 2005. Forecasting stock market movement direction with support vector machine. Comput. Oper. Res. 32 (10), 2513–2522.
Jiang, H., Shen, S., Zhang, T., 2012. Stock Market Forecasting Using Machine Learning Algorithms. Department of Electrical Engineering. CA. Stanford University, Stanford, pp. 1–5.
Kaczmarek, T., Będowska-Sójka, B., Grobelny, P., Perez, K., 2022. False safe haven assets: evidence from the target volatility strategy based on recurrent neural network. Res. Int. Bus. Financ. 60, 101610.
Kijewski, M., Ślepaczuk, R., 2020. Predicting prices of S&P500 index using classical methods and recurrent neural networks, working papers of faculty of economic. Sci., Univ. Wars., WP 27/2020, 333.
Kość, K., Sakowski, P., Ślepaczuk, R., 2019. Momentum and contrarian effects on the cryptocurrency market. Phys. A: Stat. Mech. its Appl. 523, 691–701.
Lane, G., 1984. Lane’s Stochastics. Tech. Anal. Stocks Commod. 2, 87–90.
Leigh, W., Purvis, R., Ragusa, J.M., 2002. Forecasting the NYSE composite index with technical analysis, pattern recognizer, neural network, and genetic algorithm: a case study in romantic decision support. Decis. Support Syst. 32 (4), 361–377.
Lin, C.-H., Lin, S.-Y., Wu, M.-C., 2006. An effective application of decision tree to stock trading. Expert Syst. Appl. 31 (2), 270–274.
Liu, Y., Li, Z., Nekhili, R., Sultan, J., 2023. Forecasting cryptocurrency returns with machine learning. Res. Int. Bus. Financ., 101905
Madan I., Saluja S., Zhao A., 2015. Automated Bitcoin Trading via Machine Learning Algorithms. URL: http://cs229. stanford. edu/proj2014/Isaac% 20Madan, 20.
Magdon-Ismail, M., Pratap, A., Atiya, A., Abu-Mostafa, Y., 2004. On the maximum drawdown of a brownian motion. J. Appl. Probab. 41 (1), 147–161.
Michańków, J., Sakowski, P., Ślepaczuk, R., 2022. The comparison of LSTM in algorithmic investment strategies on BTC and SP500 index. Sensors 22 (3), 917. https://doi.org/10.3390/s22030917.
Murphy, J.J., 1986. Technical analysis of the futures markets. A Comprehensive Guide to Trading Methods and Applications.
Murty, M.N.,Devi, V.S., 2011. Pattern recognition: An algorithmic approach. Springer Science & Business Media.
Nelder, J., Wedderburn, R., 1972. Generalized linear models. J. R. Stat. Soc.: Ser. A) 135 (3), 370–384.
Ren, Y.S., Ma, C.Q., Kong, X.L., Baltas, K., Zureigat, Q., 2022. Past, present, and future of the application of machine learning in cryptocurrency research. Res. Int. Bus. Financ. 63, 101799.
Sharpe, W., 1966. Mutual fund performance. J. Bus. 39 (1), 119–138.
Sortino, F.A., Price, L.N., 1994. Performance measurement in a downside risk framework. J. Invest. 3 (3), 50–58.
Tian, G., Wan, G., Guo, M., 2002. Market efficiency and the returns to simple technical trading rules: new evidence from U.S. equity market and Chinese equity markets. Asia-Pac. Financ. Mark. 9, 241–258.
Wilder J., 1978. New Concepts in Technical Trading Systems. Trend Research.
Williams L., 1979. How I Made One Million Dollars Last Year Trading Commodities. Windsor Books.
Young, T., 1991. Calmar ratio: a smoother tool. Futures 20 (1), 40.
Zenkova, M., Ślepaczuk, R., 2018. Robustness of support vector machines in algorithmic trading on cryptocurrency market Cent. Eur. Econ. J 5 (52), 186–205.