Bank i Kredyt 56(3), 2025, 339-360

BANK I KREDYT
BANK & CREDIT

Is the inclusion of a broad set of explanatory variables relevant in EPS forecasting? Evidence from Poland

Wojciech Kuryłek*

Submitted: 25 May 2023. Accepted: 22 November 2024.

DOI: 10.5604/01.3001.0055.1459

Abstract

This study explores the critical role of accurate earnings forecasts for publicly traded firms in achieving investment success, particularly in markets with limited analyst coverage, such as emerging markets like Poland. It evaluates the precision of forecasts generated by a wide array of explanatory variables, including accounting, market, and macroeconomic factors, employing gradient-boosting decision tree machine learning, multilayer perceptron networks, and convolution networks, contrasted with a seasonal random walk model. These models are applied to EPS data from companies listed on the Warsaw Stock Exchange from 2008 to 2019. Multivariate methods are trained using a comprehensive set of 1,598 explanatory variables, encompassing company-specific financial and market metrics along with macroeconomic indicators. The seasonal random walk model demonstrated the lowest error, as measured by the Mean Arctangent Absolute Percentage Error (MAAPE), findings validated through rigorous statistical examinations. Various robustness checks, employing diverse timeframes and error metrics, reaffirm this outcome. The dominance of a simplistic model may arise from the overfitting tendencies of complex models and the relatively straightforward dynamics observed in Polish listed companies.

Keywords: earnings per share, random walk, gradient-boosting decision tree, multilayer perceptron network, convolution network, Warsaw Stock Exchange

* University of Warsaw, Faculty of Management; e-mail: wkurylek@wz.uw.edu.pl; ORCID: 0000-0003-0692-3300.

1. Introduction

The pricing of company stocks relies heavily on the multiplication of earnings per share (EPS) by the price-to-earnings multiple, a critical step in investment analysis. Accurate forecasting of these components is paramount, with EPS predictions assuming particular significance. They offer essential numerical insights into a company's future trajectory, providing valuable data on potential market valuation and informing auditing expectations. While financial analysts extensively cover companies in developed markets like the US, only a fraction, approximately 20%, receive similar attention in emerging markets such as Poland. Thus, there exists a compelling need to utilize statistical or machine learning models for EPS forecasting.

This article conducts a comparative evaluation of models, employing an extensive set of explanatory fundamental, market, and macroeconomic variables, utilizing the gradient-boosting decision tree (XGBoost) machine learning method, the multilayer perceptron artificial neural network (MLP), and the convolution neural network (CNN). It encompasses quarterly EPS data for 267 companies listed on the Polish stock exchange and uses 1,598 explanatory variables from the 2008–2009 financial crisis through the 2020 pandemic.

Instead of relying solely on the conventional mean absolute percentage error (MAPE) metric, which is susceptible to extreme values when the denominator is small, an alternative measure, the mean arctangent absolute percentage error (MAAPE) proposed by Kim and Kim (2016), is calculated and utilized in this study.

In summary, this article pursues several objectives. Firstly, it aims to evaluate the performance of various cutting-edge machine learning and deep learning methods over a very wide set of explanatory variables in EPS prediction. Secondly, it seeks to apply diverse error metrics, varying timeframes, and a range of statistical tests to validate the outcomes of these experiments. Thirdly, it endeavours to adapt and employ a relative performance error metric to address scenarios where actual profits approach zero, utilizing MAAPE as an error metric. Additionally, it applies other popular error metrics as a robustness check. Lastly, it aims to elucidate the practical implications of these findings for investment strategies in Polish stocks.

2. Literature review

The forecasting of EPS through algorithms commenced in the 1960s, prompting scholarly investigation primarily focused on autoregressive integrated moving average (ARIMA) models (Ball, Watts 1972; Watts 1975; Griffin 1977; Foster 1977; Brown, Rozeff 1979). These models constituted the main category under scrutiny, with research outcomes exhibiting variability: while certain studies advocated for the simplicity of the basic random walk model, suggesting that more complex models did not consistently outperform it, others arrived at divergent conclusions. A similar inquiry into the Polish market was conducted by Kuryłek (2023a).

Over time, a consensus began to form favouring ARIMA-type models for their typically accurate forecasts (Lorek 1979; Bathke, Lorek 1984). This consensus endured until the late 1980s, when a prevailing belief arose suggesting that forecasts made by financial analysts surpassed those generated by time series models (Brown et al. 1987). However, Conroy and Harris (1987) noted the superiority of analysts' forecasts in short-term horizons, which diminished over longer periods. This perspective persisted until recent years, when the superiority of analysts over time series models was once again called into question (Lacina, Lee, Xu 2011; Bradshaw et al. 2012; Pagach, Warr 2020; Gaio et al. 2021).

Furthermore, since the late 1960s, researchers have explored various approaches employing exponential smoothing for EPS prediction (Elton, Gruber 1972; Ball, Watts 1972; Johnson, Schmitt 1974; Brooks, Buckmaster 1976; Ruland 1980; Brandon, Jarrett, Khumawala 1987; Jarrett 2008), yielding mixed findings. The research for the Polish market was conducted by Kuryłek (2023b).

Lorek and Willinger (1996) demonstrated the superiority of multivariate cross-sectional models over firm-specific and common-structure ARIMA models. Lev and Thiagarajan (1993) identified 12 fundamental signals from financial ratios, subsequently utilized by Abarbanell and Bushee (1997) for EPS forecasting. Similar fundamental variables were employed by Cao and Gan (2009), Cao and Parry (2009), Ahmadpour, Etemadi and Moshashaei (2015), and Ball and Ghysels (2017) for multivariate EPS forecasting using neural networks, confirming their effectiveness. Ohlson (1995, 2001) as well as Olhson and Juettner-Nauroth (2005) formulated a residual earnings model, while Pope and Wang (2005, 2014) established theoretical frameworks linking earnings forecasts to accounting variables and stock prices. Li (2011) developed a model for forecasting earnings for loss-making firms, demonstrating its efficacy. Lev and Souginannis (2010) provided evidence of the usefulness of estimate-based accounting items for predicting next year's earnings, albeit with limited success in subsequent years. Hou, van Dijk and Zhang (2012) achieved substantial R2 coefficients in cross-sectional regression models for earnings forecasting. Li and Mohanram (2014) compared various models, revealing the superiority of some over others. Harris and Wang (2019) found Pope and Wang's (2005) model generally less biased and more accurate.

Recent studies have highlighted the importance of artificial neural networks in predicting EPS, yielding mixed outcomes. Atiya, Shaheen, and Talaat (1997) demonstrated the effectiveness of neural networks based on fundamental characteristics for stock price prediction, consistently outperforming other methods. Conversely, Lai and Li (2006) found that an ANN model had the lowest accuracy in EPS prediction. Cao, Schniederjans and Zhang (2004) compared neural feedforward networks (MLP) to other models, showing their superior accuracy. Cao and Parry (2009) consistently favoured univariate neural network models over linear regression, revealing that genetic algorithms improved performance in estimating network weights. Similarly, Cao and Gan (2009) confirmed the superiority of neural networks, particularly when optimized using genetic algorithms, for predicting EPS in Chinese companies. Gupta, Khirbat, and Singh (2013) identified an optimal architecture for stock market forecasting using multilayer perceptron networks, emphasizing the role of factors like EPS and public confidence. Ahmadpour, Etemadi, and Moshashaei (2015) achieved remarkable success with standard MLP neural networks, with extracted rules showing higher accuracy than pure MLP models. Chen et al. (2020) explored various prediction methods, with ensemble methods proving the most accurate. Elend et al. (2020) found that LSTM networks outperformed other models in predicting EPS. Suler, Vochozka, and Vrbka (2020) successfully used LSTM networks for bankruptcy prediction in the Czech Republic. Decision tree-based techniques for EPS forecasting yielded mixed results (Delen, Kuzey, Uyar 2013; Gramacy, Gerakos 2013; Elamir 2020; Chen et al. 2020). Recent advances in machine learning and deep learning facilitated innovative experiments. Xiaoqiang (2022) provided an overview of machine learning techniques for financial ratios forecasting, including EPS. Furthermore, Xiaoqiang's (2022) article covered deep learning and machine learning techniques like convolutional neural networks and decision trees for EPS prediction. In the latest research, Dreher, Eichfelder, and Noth (2024) found that considering tax loss carryforwards did not improve EPS forecasts for German listed companies and sometimes worsened predictions in out-of-sample tests, using tax footnotes information.

The incorporation of macroeconomic and market variables in EPS forecasting has existed for nearly three decades. Pioneering research by Chant (1980) employed three macroeconomic variables: growth in money supply, bank loans, and the stock market index, to predict corporate earnings. His findings suggested that models utilizing these leading economic indicators yielded lower errors compared to pure time series models. Lev (1980) observed that simple-index models utilizing gross national product (GNP) and total corporate profits after taxes as dependent variables for predicting sales, operating income, and net income generally outperformed random walk models. Lev and Thiagarajan (1993), in a seminal work, proposed that macroeconomic variables such as annual changes in CPI, GNP, and business inventories could significantly influence corporate performance.

Bansal, Strauss, and Nasseh (2015) conducted simulations to generate out-of-sample forecasts for a large set (21) of individual autoregressive distributed lag (ARDL) models. Each model incorporated one plausible predictor variable identified from prior research. These variables encompassed not only firm-specific accounting data, but also market-wide and macroeconomic factors, including S&P 500 dividend yield, S&P 500 PE ratio, total earnings for the S&P, Dow Jones dividend yield, corporate bond yield for long-term AAA-rated corporate bonds, default spread between BAA- and AAA-rated corporate bonds, yield spread (difference between 10-year Treasury bond and 3-month Treasury bill yields), Treasury bill yield (3-month), and CPI inflation.

Ball and Ghysels (2017) employed an ARDL framework that incorporated two firm-level stock market predictor variables: excess stock returns and return volatility. Additionally, their model included the following macroeconomic predictor variables: industrial production, CPI inflation index, oil price growth, 3-month Treasury bill yields, and bond term spread (10-year Treasury vs. 3-month Treasury bill), default spread (BAA vs. AAA corporate bonds).

3. Data and methodology

3.1. Data

The Polish stock market, integrated into the European Union post-2004, boasts considerable depth, with a market capitalization reaching USD 197 billion and accommodating 774 listed companies by the end of 2021. Despite this, analyst coverage for these stocks is notably limited compared to the United States or Western Europe, with only around 20% of the 711 listed companies receiving attention in 2019. This underscores the need for statistical forecasting of critical financial data using analytical methodologies. This study primarily focuses on the earnings per share (EPS) data series and other financial explanatory variables sourced from EquityRT, a financial analysis platform. EPS patterns of companies listed on the Warsaw Stock Exchange are analysed from Q1 2010 to Q4 2019, covering significant structural shifts like the 2008–2009 financial crisis and the onset of the COVID-19 pandemic in 2020. Since 2020, a period of turbulence began, initiated by the COVID-19 pandemic in 2020, followed by the war in Ukraine in 2022, which triggered the energy crisis. Therefore, the year 2019 was the last year of relative calm, a period in which models could be tested. If a model is unable to achieve good predictive results during a period of stability and predictability, it is even less capable during periods of market disruption, when extrapolating trends from the past into the future (as all models do) is inherently fraught with significant error. Furthermore, the author's intention was to maintain the same sample as in other articles on the same research topic (Kuryłek 2023a, 2023b, 2024) to ensure precise comparability of results. For forecasting, data from Q1 2010 to Q4 2018 (36 quarters) are utilized for model estimations, with Q1 2019 to Q4 2019 data reserved for out-of-time validation testing. Forecast horizons range from 1 to 4 quarters ahead, with additional validation samples from the years 2017 and 2018. The dataset, after comprehensive coverage and excluding splits and reverse splits, comprises 267 companies. Also companies that ceased publishing financial reports during the study period were excluded from the sample.

The explanatory variables

- The seasonal random walk model (SRW)

  The SRW can be described as:

$$EPS_t = EPS_{t-4} + \varepsilon_t \text{ where } \varepsilon_t \text{ are IID and } \varepsilon_t \sim N\left(0, \sigma^2\right) \tag{1}$$

The forecast represented by, $\widehat{EPS}_t = EPS_{t-4}$ utilizes the value delayed by 4 quarters as the prediction, eliminating the need for parameter estimation. This model serves as a benchmark, as evidenced by Kuryłek (2023a, 2023b), showcasing its superiority over time series models specifically within the Polish context.

- The multivariate model with all available variables (ALL)

  Due to the challenge of identifying crucial variables in EPS forecasting, the author opts to incorporate nearly all variables accessible within the EquityRT platform. Consequently, the multivariate model can be expressed as follows:

$$EPS_{t+4} = f\left(EPS_t, EPS_{t-1}, EPS_{t-8}, X_t^1, .., X_t^n, X_{t-1}^1, .., X_{t-1}^n, X_{t-4}^1, .., X_{t-4}^n, Y_t^1, Y_t^2\right) + \varepsilon_t \tag{2}$$

Considering the need to capture both year-to-year and quarter-to-quarter dynamics, the model incorporates lags of the independent variables at 4, 5, and 8 quarters. The exclusion of lags $t-1$, $t-2$, and $t-3$ stems from the assumption that a single model will be used to forecast one year ahead, rather than employing four separate models for the first, second, third, and fourth quarters. Therefore, to generate a forecast for the fourth quarter of 2019, for instance, explanatory variables from the first, second, and third quarters are not required, as these data points are not yet available before the start of 2019. Thus, the independent variables (designated as X_t^i and Y_t^i) encompass factors lagged by these periods, including the lagged dependent variable, 67 accounting fields reported from financial statements on a per-share basis, 99 financial ratios, and 3 firm-specific market variables (end-of-quarter share price, mean price within the quarter, and price standard deviation). Additionally, 362 macro variables are included, sourced from various databases contained in the EquityRT such as the International Monetary Fund (189 variables), OECD (143 variables), Eurostat (56 variables), and the Bank of International Settlements (20 variables). Furthermore, 2 industry classification variables (Y_t^1 and Y_t^2) available on the EquityRT platform are incorporated without delay, as they may play a role as potentially explanatory variables. Consequently, the vector of independent variables spans 1,598 dimensions, and the model is trained using 7,476 observations (28 quarters multiplied by 267 companies) to predict EPS for the year 2019.

3.2. Estimation techniques

The XGBoost (XGB) – a gradient-boosting decision tree

XGBoost, an abbreviation for eXtreme Gradient Boosting, was initially introduced by Chen and Guestrin (2016) as a significant improvement over the original gradient boosting algorithm, renowned for its speed and effectiveness. This advanced machine learning technique, rooted in decision tree algorithms from the 1960s, is widely employed in regression and classification tasks. It constructs a prediction model by combining weak prediction models, typically simple decision trees, in an ensemble fashion. Each iteration of trees aims to rectify errors made by the previous ones, with the gradient descent algorithm iteratively adjusting the weights of these weak learners. This iterative process continues until the loss function is minimized or a predefined stopping criterion is met. XGBoost incorporates various techniques to enhance the performance of the gradient boosting model, including regularization to mitigate overfitting by imposing penalties on the loss function, tree pruning to remove redundant branches and improve model stability, and parallelization to expedite the training process. Furthermore, XGBoost adeptly captures inherent nonlinearities in the data. For a deeper understanding, Simon's (2020) book offers valuable insights. It would be ideal to minimize a loss function (objective function or internal metric) that is equivalent to the external metric, i.e. MAAPE. According to the requirements of the xgb library, implementing a custom loss function requires calculating its first derivative (gradient) and second derivative (Hessian). The problem is that MAAPE is not a function with a first and second derivative over its entire domain. Therefore, the most standard loss function, reg: squarederror, i.e. mean square error (MSE), was used. This methodology has been implemented using the xgb library in Python, with hyperparameters fine-tuned (Banerjee 2020) by the hyperopt library to optimize forecast performance. Optimized were the following hyperparameters: colsample_bytree, learning_rate, max_depth, min_child, n_estimators, reg_alpha, reg_lambda, and subsample. The optimal choice of them was: colsample_bytree (0.8659879434985166), learning_rate (0.13438562077762467), max_depth (5), min_child_weight (9), n_estimators (55), random_state (350), reg_alpha: (1.0070608627374555), reg_lambda (0.7810551944833559), subsample: (0.9351229731753115). Ultimately, 31 variables were selected out of 1,598 to construct the boosting trees, and the list of these variables is provided in the Appendix.

Artificial neuron networks (ANNs)

The artificial neural networks examined in this research utilized the TensorFlow module in Python for training. These networks are characterized as feedforward, indicating that data progresses unidirectionally from the input layer to the output layer. Artificial neural networks (ANNs) are commonly applied to investigate cause-effect relationships in intricate systems, often within the framework of large datasets. However, they can also be employed with smaller datasets, as demonstrated in fields such as health sciences by Pasini (2015), similar to the context of this study. Hyperparameters, including network width and depth (i.e. the number of neurons per layer and the number of layers), were optimized using the hyperas library in Python. The models underwent training via the backpropagation algorithm based on gradient descent, with only 20 epochs utilized (where one epoch constitutes a complete run-through of the training set). Backpropagation, introduced by Werbos (1988) in the late 1980s, is a standard technique for training neural networks, involving the backward propagation of errors. It adjusts the weights of a neural network based on the error rate from the previous epoch, aiming to minimize the error by feeding it back through the network. The learning rate dictates the pace of adjustment, with the objective of minimizing error rates and improving the model's generalization capability. Once a certain number of epochs are completed, the algorithm converges to a state where there is minimal change in loss over subsequent epochs, typically reaching a local optimum of the defined loss function. Input parameters are often standardized for ANNs when dealing with multivariate data. In all models analysed, the hyperbolic tangent (tanh) activation function, a popular choice, was applied across all layers. Additionally, the weights between layers were initialized using the glorot_uniform initializer proposed by Bengio and Glorot (2010), which generates initial weight values from a uniform distribution. For further insights into various network architectures and parameters, refer to the book by Bengio, Courville, and Goodfellow (2017). The input data for the neural networks were normalized using the MinMaxScaler() function from the sklearn library. Due to the large number of explanatory variables from multiple sources and their potential multicollinearity, the system could exhibit numerical instability; nevertheless, it converged to a minimum in practice in all analysed cases.

The multilayer perceptron network (MLP)

An artificial neural network known as a multilayer perceptron (MLP) comprises multiple layers of interconnected nodes. Each layer's nodes establish connections with those in the subsequent layer, and the connection weights are learned during training. Typically, an MLP consists of three or more layers: an input layer, one or more hidden layers, and an output layer. Within each hidden layer, the output of each node results from a weighted sum of the preceding layer's node outputs, augmented by a bias term. MLP neural networks trace their origins back to 1958, when Rosenblatt (1958) introduced a layered network of perceptrons, inspired by the brain's functionality. The number of layers and neurons constitute the network's hyperparameters, which require fine-tuning. While deeper neural networks excel in data processing, excessively deep layers can pose challenges such as vanishing gradients and overfitting. Empirical rules of thumb guide the determination of the optimal number and size of hidden layers, as detailed in Heaton's (2008) book. According to this source, a single hidden layer is sufficient to approximate any function. Therefore, the network in this study was designed with one hidden layer. Additionally, a widely endorsed guideline suggests that the optimal size of the hidden layer should lie approximately between that of the input and output layers. In this instance, the hidden layer's size corresponds to the mean of the sizes of the input and output layers.

The convolution neural network (CNN)

A convolutional neural network (CNN) is a specialized type of feed-forward neural network designed to process data with a grid-like topology, such as images. CNNs draw inspiration from the functioning of the visual cortex in animals. The fundamental component of a CNN is the convolution layer, which takes an input image and applies a filter to extract features. This filter, also known as a kernel, is a small matrix of weights that moves across the input image, producing a feature map that highlights important image characteristics. Through the optimization of filters, the network autonomously learns feature engineering. Additionally, CNNs often incorporate pooling layers to reduce the size of feature maps, thereby enhancing computational efficiency by summarizing information in each region. Following pooling, the image is flattened into a one-dimensional vector, and a dense layer from the MLP network with one input and output layer is applied. The architecture of CNNs traces its origins to Fukushima's "neocognitron" introduced in 1980, with modern CNNs evolving in the 1990s, building upon this foundation. To mitigate issues like vanishing and exploding gradients observed in earlier neural networks, CNNs utilize regularized weights over fewer connections (Fukushima 1980). Recently, CNNs have found applications in financial time series analysis, such as stock price prediction by Gabbouj et al. (2017) and EPS forecasting by Elend et al. (2020). In the analysis presented here, the input is a vector of length 1,598, which can be conceptualized as a rudimentary image of dimensions 34 × 47. An illustration of such an image is depicted in Figure 1.

3.3. Mean Arctangent Absolute Percentage Error (MAAPE)

Using the symbols $A_1^i, \ldots, A_4^i$, to represent the actual earnings per share (EPS) from the first to the fourth quarter of 2019 for a specific firm *I*, and $F_1^i, \ldots, F_4^i$ as the corresponding forecasts (i.e. $\hat{Q}_t$, where t = 37,.., 40 for *i*-th company), the absolute percentage error (APE) of such prediction during the *j*-th quarter of 2019, for any firm *i*, can be formulated as follows:

$$APE_j^i = \left| \frac{A_j^i - F_j^i}{A_j^i} \right| \tag{3}$$

Nonetheless, the absolute percentage error (APE) presents a notable constraint: it can yield infinite or undefined values as the actual figures approach or reach zero, a scenario often encountered in earnings forecasts. Additionally, exceedingly low actual figures, usually below one, can lead to significant percentage errors (outliers). Moreover, when actual values are zero, the APE becomes infinite or undefined. To mitigate this challenge, Kim and Kim (2016) introduced the arctangent absolute percentage error as an innovative solution in this field.

$$AAPE_j^i = \begin{cases} 0 & \text{if } A_j^i = F_j^i = 0 \\ arctan\left(\left| \frac{A_j^i - F_j^i}{A_j^i} \right| \right) & \text{otherwise} \end{cases} \tag{4}$$

The rationale behind this approach arises from the nature of the arctan function, which transforms values spanning from -∞ to +∞ into the range of [-π/2, π/2]. Consequently, the MAAPE for the *j*-th quarter across all *I* companies in the dataset can be formulated as:

$$MAAPE_j = \frac{1}{I}\sum_{i=1}^{I} AAPE_j^i = \frac{1}{I}\sum_{i=1}^{I} arctan\left(\left|\frac{A_j^i - F_j^i}{A_j^i}\right|\right) \tag{5}$$

The selection of MAAPE over MAPE (Mean Absolute Percentage Error) was deliberate due to the presence of companies with actual profits nearing zero within the examined dataset. In scenarios where only a single observation approaches zero while others remain significantly distant, the MAPE for that specific observation can escalate to an extremely large value, approaching infinity. This phenomenon could potentially overshadow the mean calculation, rendering the remaining observations inconsequential.

3.4. The statistical test

To evaluate the statistical significance of variations in MAAPE among multiple models, a nonparametric two-sided Wilcoxon test, as elucidated by Wilcoxon (1945), is utilized. This test functions as a paired difference test for two matched samples. It is noteworthy that this test doesn't necessitate specific assumptions regarding a probability distribution, except for the symmetry of the difference in scores and the independence of these differences. Ruland (1980) provided detailed insights into the implementation of the Wilcoxon test in validation, particularly in determining whether errors generated by different EPS models exhibit statistical differences. Distinct tables containing p-values are generated for each quarter, spanning from one to four, as well as for all quarters collectively.

$$H_0\text{: medians of AAPEs of a pair of models are the same} \tag{6}$$

Should the p-values of each test drop below the predetermined significance threshold of 0.05, the null hypothesis for each test would be deemed invalid. This principle, commonly employed, draws from multiple sources, such as Ruland (1980).

4. Results

4.1. Empirical findings

The seasonal random walk (SRW) model, as delineated in Table 1, consistently outperforms all other models throughout each quarter and in the entirety of 2019, demonstrating superior overall performance and registering the lowest MAAPE. Conversely, the multilayer perceptron model (MLP) displayed the weakest performance among all the approaches. Following closely behind the best was the XGBoost (XGB) approach, while the convolution neural network (CNN) fared slightly worse.

To ascertain whether the errors of the top-performing model significantly differ from those of the other approaches, the Wilcoxon nonparametric test was employed to compare the AAPE medians between the SRW model and all other methods. As illustrated in Table 2, the findings indicate that the seasonal random walk (SRW) model consistently showcases statistically lower errors compared to the other approaches across all the analysed periods, except for the 2nd quarter of 2019. In this quarter, the results of the seasonal random walk and the XGBoost are not statistically distinguishable at the 0.05 statistical significance level. Furthermore, in the 4th quarter, the results of the Wilcoxon test for the seasonal random walk and XGBoost combination, as well as the seasonal random walk and the convolution neural network, only marginally fall below this threshold.

4.2. Robustness checks

Robustness evaluations were conducted across different years and two popular error metrics. It is noteworthy that throughout all the examined years – 2017, 2018, and 2019 – the seasonal random walk model consistently yielded superior forecasts compared to alternative methods, as illustrated in Table 3. In 2018, the least effective approach was the XGBoost, while in 2017, it was the convolution neural network. Surprisingly, the performance of the multilayer perceptron network (MPL) exhibited a remarkable improvement in these years compared to 2019. Additionally, the Wilcoxon test was utilized to compare all model pairs against the seasonal random walk model, and the corresponding p-values for each year are detailed in Table 4. Across all these years, the seasonal random walk model consistently demonstrated statistically superior outcomes compared to alternative methods. Therefore, the continual dominance of the seasonal random walk model becomes evident over time.

Table 5 presents an evaluation of the performance of the analysed models using alternative error metrics: Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). This evaluation covers all quarters aggregated for the year 2019. To ensure a fair comparison, these metrics were adjusted for Consumer Price Index (CPI) inflation. This adjustment guarantees equivalence in the present value of future errors in nominal terms with current errors. Consistent with previous observations in 2019, the seasonal random walk model exhibited the lowest errors across all metrics, including both RMSE and MAE. Conversely, the multilayer perceptron network performed the worst in 2019 according to both error metrics, as seen previously in the case of MAAPE.

The p-values contained in Table 6, based on the Wilcoxon test, underscore significant differences between the outcomes of the SRW model and other model pairings. However, for the combination of XGBoost and the seasonal random walk, the p-value is marginally below the 0.05 significance level, specifically concerning the RMSE metric.

The research findings indicate that despite incorporating an extensive array of explanatory variables encompassing financial, market, and macroeconomic factors, and utilizing advanced forecasting algorithms based on them, there is no improvement in performance. Furthermore, these sophisticated approaches fail to outperform the basic seasonal random walk method.

4.3. Discussion

The relatively inferior performance of more complex models, such as those relying on gradient-boosting decision trees or artificial neural networks, can be attributed to overfitting. Overfitting leads to unstable relationships among variables, contingent on the specific test dataset. Utilizing such relationships for predictions is only reasonable if the statistical relationship is sufficiently robust, as highlighted by Lev and Souginannis (2010). This observation is consistent with the findings of Dreher, Eichfelder and Noth (2024), who also demonstrated, for German listed companies, that complex deep learning approaches, while optimizing explanatory power within the sample, do not perform well for out-of--sample prediction. These sophisticated models risk overparameterizing the market's straightforward behaviour, resulting in larger forecast errors.

The rationale behind the superior performance of simpler models may align with the Polish scenario. Advanced models often tend to be overly intricate, possessing an excess of parameters to describe relatively straightforward economic phenomena. This observation further corroborates the research conducted by Kuryłek (2023a, 2023b), which showed that even basic models like ARIMA and exponential smoothing, effective for the US market, were outperformed by the simple seasonal random walk model in Poland. This reinforces the hypothesis that the inherent simplicity of the Polish stock market likely underpins the effectiveness of the seasonal random walk model, or alternatively, additional calibration for out-of-sample predictions might be necessary.

Hence, straightforwardly applying any of these sophisticated techniques beyond the conventional seasonal random walk in Poland for EPS forecasting in investment contexts appears impractical. Furthermore, considering that EPS behaviour follows a seasonal random walk and acknowledging that stock prices are derived from the multiplication of the P/E multiple by EPS, one might infer that stock prices exhibit at least as much randomness as EPS. Since EPS behaviour, characterized by a random walk, is inherently challenging, accurately predicting stock prices for a period extending at least one quarter ahead becomes even more daunting.

In shorter timeframes, when EPS remains constant, stock price forecasting behaves similarly to P/E multiples. Consequently, exploring methods to forecast P/E multiples for periods shorter than one quarter, occurring between the publication of quarterly financial reports, could be of significant interest from an investment perspective. The forecast generated by the seasonal random walk essentially represents a value from the corresponding quarter of the previous year. This implies that for predicting future prices, even over extended horizons, the P/E multiple might carry more significance than next year's earnings of companies.

This aligns with economic theory, which suggests that the P/E multiple is influenced by expected future earnings growth, future interest rates, and market sentiment or premium reflecting investor risk appetite, whereas EPS forecasts pertain only to near-future earnings. In both short-term and long-term contexts, the conclusion is clear: for investment, the P/E multiple holds greater importance than EPS prediction.

5. Conclusions

The study investigates the predictive capacities of four methodologies: the seasonal random walk (SRW), gradient boosting tree (XGB), multilayer perceptron network (MLP), and convolutional neural network (CNN). These multivariate approaches are trained using a very comprehensive set of 1,598 explanatory variables, encompassing firm-specific financial and market variables alongside macroeconomic indicators. EPS forecasting holds significant importance in emerging markets like Poland, where financial analyst coverage of listed companies is limited. Analysing quarterly EPS data from 267 Polish firms over the period 2010 to 2019, the SRW model consistently exhibited the lowest error rates, offering a more precise depiction of the Polish market compared to alternative models. Moreover, the SRW model consistently outperformed other methods across various periods and error metrics such as RMSE or MAE. This trend is supported by Wilcoxon tests and can be attributed to the overparameterization of complex models, their proneness to overfitting, and the relatively simplistic nature of the Polish stock market.

The research findings suggest that despite the inclusion of a wide range of explanatory variables covering financial, market, and macroeconomic factors, and the application of advanced forecasting algorithms based on them, there is no enhancement in performance.

The practical implication of this research suggests that employing techniques beyond the standard seasonal random walk for EPS forecasting in Poland lacks practical justification. However, depending on the seasonal random walk for EPS modelling implies that forecasted stock prices may exhibit significant randomness, posing challenges for prediction. Therefore, forecasting the P/E multiple might hold more significance than predicting EPS for future stock price projections, particularly in shorter investment horizons when EPS remains constant.

Future research avenues could explore the correlation between forecasting accuracy and firm size, with industry sector analysis potentially influencing the selection of the most suitable EPS forecasting model. Investigating time series transformations to normalize EPS distributions could provide valuable insights. Additionally, incorporating analysis of the text from companies' public communications presents an intriguing prospect. Comparing the predictive accuracy of the best algorithmic model with forecasts from market analysts offers another interesting avenue. Furthermore, assessing the performance and accuracy of various predictive models and financial analysts' projections during economic downturns, such as the 2008–2009 financial crisis or the COVID-19 pandemic, and the onset of the war in Ukraine, could yield valuable insights. Identifying seasonal patterns through the SRW model may provide insights into investment strategies, potentially challenging the "weak form" of the Efficient Market Hypothesis (EMH). Further research could also consider dividing the analysed period into sub-periods of bull and bear markets (in reference to the WIG index).

References

Abarbanell J., Bushee B. (1997), Fundamental analysis, future EPS, and stock prices, *Journal of Accounting Research*, 35(1), 1–24, DOI: 10.2307/2491464.

Ahmadpour A., Etemadi H., Moshashaei S. (2015), Earnings per share forecast using extracted rules from trained neural network by genetic algorithm, *Computational Economics*, 46(1), 55–63, DOI: 10.1007/s10614-014-9455-6.

Atiya A., Shaheen S., Talaat N. (1997), *An efficient stock market forecasting model using neural networks*, IEEE International Conference on Neural Networks, DOI: 10.1109/ICNN.1997.614231.

Ball R., Ghysels E. (2017), Automated earnings forecasts: beat analysts or combine and conquer?, *Management Science*, 64(10), 4936–4952, DOI: 10.1287/mnsc.2017.2864.

Ball R., Watts R. (1972), Some time series properties of accounting income, *The Journal of Finance*, 27(3), 663–681, DOI: 10.1111/j.1540-6261.1972.tb00991.x.

Banerjee P. (2020), *A guide on XGBoost hyperparameters tuning*, https://www.kaggle.com/code/prashant111/a-guide-on-xgboost-hyperparameters-tuning.

Bansal N., Nasseh A., Strauss J. (2015), Can we consistently forecast a firm's earnings? Using combination forecast methods to predict the EPS of Dow firms, *Journal of Economics and Finance*, 39(1), 1–22, DOI: 10.1007/s12197-012-9234-y.

Bathke Jr. A.W., Lorek K.S. (1984), The relationship between time-series models and the security market's expectation of quarterly earnings, *The Accounting Review*, 59(2), 163–176.

Bradshaw M., Drake M., Myers J., Myers L. (2012), A re-examination of analysts' superiority over time-series forecasts of annual earnings, *Review of Accounting Studies*, 17(4), 944–968, DOI: 10.1007/s11142-012-9185-8.

Bengio Y., Courville A., Goodfellow I. (2017), *Deep Learning*, The MIT Press.

Bengio Y., Glorot X. (2010), *Understanding the difficulty of training deep feedforward neural networks*, 13th International Conference on Artificial Intelligence and Statistics, 9, 249–256.

Brandon Ch., Jarrett J.E., Khumawala S.B. (1987), A comparative study of the forecasting accuracy of Holt-Winters and economic indicator models of earnings per share for financial decision making, *Managerial Finance*, 13(2), 10–15, DOI: 10.1108/eb013581.

Brooks L.D., Buckmaster D.A. (1976), Further evidence of the time series properties of accounting income, *The Journal of Finance*, 31(5), 1359–1373, DOI: 10.1111/j.1540-6261.1976.tb03218.x.

Brown L.D., Griffin P.A., Hagerman R.L., Zmijewski M.E. (1987), Security analyst superiority relative to univariate time-series models in forecasting quarterly earnings, *Journal of Accounting and Economics*, 9(1), 61–87, DOI: 10.1016/0165-4101(87)90017-6.

Brown L.D., Rozeff M.S. (1979), Univariate time-series models of quarterly accounting earnings per share: a proposed model, *Journal of Accounting Research*, 17(1), 179–189, DOI: 10.2307/2490312.

Cao Q., Gan Q. (2009), *Forecasting EPS of Chinese listed companies using a neural network with genetic algorithm*, 15th Americas Conference on Information Systems, AMCIS 2009, 2791–2981.

Cao Q., Parry M. (2009), Neural network earnings per share forecasting models: a comparison of backward propagation and the genetic algorithm, *Decision Support Systems*, 47(1), 32–41, DOI: 10.1016/j.dss.2008.12.011.

Cao Q., Schniederjans M.J., Zhang W. (2004), Neural network earnings per share forecasting models: a comparative analysis of alternative methods, *Decision Sciences*, 35(2), 205–237, DOI: 10.1111/j.00117315.2004.02674.x.

Chant P.D. (1980), On the predictability of corporate earnings per share behavior, *The Journal of Finance*, 35(1), 13–21, DOI: 10.2307/2327177.

Chen Y., Chen S., Huang H., Sangaiah A. (2020), Applied identification of industry data science using an advanced multi-componential discretization model, *Symmetry*, 12(10), 1–28, DOI: 10.3390/sym12101620.

Chen T., Guestrin C. (2016), *XGBoost: a scalable tree boosting system*, 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794, DOI: 10.1145/2939672.2939785.

Conroy R., Harris R. (1987), Consensus forecasts of corporate earnings: analysts' forecasts and time series methods, *Management Science*, 33(6), 725–738, DOI: 10.1287/mnsc.33.6.725.

Delen D., Kuzey C., Uyar A. (2013), Measuring firm performance using financial ratios: a decision tree approach, *Expert Systems with Applications*, 40(10), 3970–3983, DOI: 10.1016/j.eswa.2013.01.012.

Dreher S., Eichfelder S., Noth F. (2024), Does IFRS information on tax loss carryforwards and negative performance improve predictions of earnings and cash flows?, *Journal of Business Economics*, 94(1), 1–39, DOI: 10.1007/s11573-023-01147-7.

Elamir E. (2020), Modeling and predicting earnings per share via regression tree approaches in banking sector: Middle East and North African countries case, *Investment Management and Financial Innovations*, 17(2), 51–68, DOI: 10.21511/imfi.17(2).2020.05.

Elend L., Kramer O., Lopatta K., Tideman S. (2020), *Earnings prediction with deep learning*, KI 2020: Advances in Artificial Intelligence, 267–274, DOI: 10.1007/978-3-030-58285-2_22.

Elton E.J., Gruber M.J. (1972), Earnings estimates and the accuracy of expectational data, *Management Science*, 18(8), B409–B424, DOI: 10.1287/mnsc.18.8.B409.

Foster G. (1977), Quarterly accounting data: time-series properties and predictive-ability results, *The Accounting Review*, 52(1), 1–21.

Fukushima K. (1980), Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position, *Biological Cybernetics*, 36(4), 193–202, DOI: 10.1007/BF00344251.

Gabbouj M., Iosifidis A., Kanniainen J., Passalis N., Tefas A., Tsantekidis A. (2017), *Forecasting stock prices from the limit order book using convolutional neural networks*, IEEE 19th conference on business informatics (CBI), DOI: 10.1109/cbi.2017.23.

Gaio L., Gatsios R., Lima F., Piamenta Jr. T. (2021), Re-examining analyst superiority in forecasting results of publicly-traded Brazilian companies, *Revista de Administracao Mackenzie*, 22(1), eRAMF210164, DOI: 10.1590/1678-6971/eramf210164.

Gramacy R.B., Gerakos J. (2013), *Regression-based earnings forecasts*, Chicago Booth Research Paper, 12–26, DOI: 10.2139/ssrn.2112137.

Griffin P. (1977), The time-series behavior of quarterly earnings: preliminary evidence, *Journal of Accounting Research*, 15(1), 71–83, DOI: 10.2307/2490556.

Gupta R., Khirbat G., Singh S. (2013), *Optimal neural network architecture for stock market forecasting,* 2013 International Conference on Communication Systems and Network Technologies, 557–561, DOI: 10.1109/csnt.2013.120.

Harris R.D.F., Wang P. (2019), Model-based earnings forecasts vs. financial analysts' earnings forecasts, *British Accounting Review*, 51(4), 424–437, DOI: 10.1016/j.bar.2018.10.002.

Heaton J. (2008), *Introduction to Neural Networks for Java*, Heaton Research Inc.

Hou K., van Dijk M., Zhang Y. (2012), The implied cost of capital: a new approach, *Journal of Accounting and Economics*, 53(3), 504–526, DOI: 10.1016/j.jacceco.2011.12.001.

Jarrett J.E. (2008), Evaluating methods for forecasting earnings per share, *Managerial Finance*, 16, 30–35, DOI: 10.1108/eb013647.

Johnson T.E., Schmitt T.G. (1974), Effectiveness of earnings per share forecasts, *Financial Management*, 3(2), 64–72, DOI: 10.2307/3665292.

Kim S., Kim H. (2016), A new metric of absolute percentage error for intermittent demand forecasts, *International Journal of Forecasting*, 32(3), 669–679, DOI: 10.1016/j.ijforecast.2015.12.003.

Kuryłek W. (2023a), The modeling of earnings per share of Polish companies for the post-financial crisis period using random walk and ARIMA models, *Journal of Banking and Financial Economics*, 1(19), 26–43, DOI: 10.7172/2353-6845.jbfe.2023.1.2.

Kuryłek W. (2023b), Can exponential smoothing do better than seasonal random walk for earnings per share forecasting in Poland?, *Bank i Kredyt*, 54(6), 651–672.

Kuryłek W. (2024), Can we profit from BigTechs' time series models in predicting earnings per share? Evidence from Poland, *Data Science in Finance and Economics*, 4(2), 218–235, DOI: 10.3934 DSFE.2024008.

Lacina M., Lee B., Xu R. (2011), An evaluation of financial analysts and naïve methods in forecasting long-term earnings, in: K.D. Lawrence, R.K. Klimberg (eds.), *Advances in Business and Management Forecasting*, Emerald Publishing, DOI: 10.1108/S1477-4070(2011)0000008009.

Lai S., Li H. (2006), The predictive power of quarterly earnings per share based on time series and artificial intelligence model, *Applied Financial Economics*, 16(18), 1375–1388, DOI: 10.1080/09603100600592752.

Lev B. (1980), On the use of index models in analytical reviews by auditors, *Journal of Accounting Research*, 18(2), 524–550, DOI: 10.2307/2490591.

Lev B., Souginannis T. (2010), The usefulness of accounting estimates for predicting cash flows and earnings, *Review of Accounting Studies*, 15(4), 779–807, DOI: 10.1007/s11142-009-9107-6.

Lev B., Thiagarajan S. (1993), Fundamental information analysis, *Journal of Accounting Research*, 31(2), 190–215, DOI: 10.2307/2491270.

Li K.K. (2011), How well do investors understand loss persistence?, *Review of Accounting Studies*, 16(3), 630–667, DOI: 10.1007/s11142-011-9157-4.

Li K.K., Mohanram P. (2014), Evaluating cross-sectional forecasting models for the implied cost of capital, *Review of Accounting Studies*, 19(3), 1152–1185, DOI: 10.1007/s11142-014-9282-y.

Lorek K.S. (1979), Predicting annual net earnings with quarterly earnings time-series models, *Journal of Accounting Research*, 17(1), 190–204, DOI: 10.2307/2490313.

Lorek K.S, Willinger G.L. (1996), A multivariate time-series model for cash-flow data, *Accounting Review*, 71, 81–101.

Ohlson J.A. (1995), Earnings, book values, and dividends in equity valuation, *Contemporary Accounting Research*, 11(2), 661–687, DOI: 10.1092/7tpj-rxqn-tqc7-ffae.

Ohlson J.A. (2001), Earnings, book values, and dividends in equity valuation: an empirical perspective, *Contemporary Accounting Research*, 18(1), 107–120, DOI: 10.1092/7tpj-rxqn-tqc7-ffae.

Ohlson J.A., Juettner-Nauroth B.E. (2005), Expected EPS and EPS growth as determinants of value, *Review of Accounting Studies*, 10(2–3), 349–365, DOI: 10.1007/s11142-005-1535-3.

Pagach D.P., Warr R.S. (2020), Analysts versus time-series forecasts of quarterly earnings: a maintained hypothesis revisited, *Advances in Accounting*, 51, 1–15, DOI: 10.1016/j.adiac.2020.100497.

Pasini A. (2015), Artificial neural networks for small dataset analysis, *Journal of Thoracic Disease*, 7(5), 953–960.

Pope P.F., Wang P. (2005), Earnings components, accounting bias and equity valuation, *Review of Accounting Studies*, 10(4), 387–407, DOI: 10.1007/s11142-005-4207-4.

Pope P., Wang P. (2014), On the relevance of earnings components: valuation and forecasting links, *Review of Quantitative Finance and Accounting*, 42, 399–413, DOI: 10.1007/s11156-013-0347-y.

Rosenblatt F. (1958), The perceptron: a probabilistic model for information storage and organization in the brain, *Psychological Review*, 65(6), 386–408, DOI: 10.1037/h0042519.

Ruland W. (1980), On the choice of simple extrapolative model forecasts of annual earnings, *Financial Management*, 9(2), 30–37, DOI: 10.2307/3665165.

Simon J. (2020), *Learn Amazon SageMaker: A guide to building, training, and deploying machine learning models for developers and data scientists*, Packt.

Suler P., Vochozka M., Vrbka J. (2020), Bankruptcy or success? The effective prediction of a company's financial development using LSTM, *Sustainability*, 12(18), 2299–2314, DOI: 10.3390/su12187529.

Watts R.L. (1975), *The time series behavior of quarterly earnings*, Working Paper, April, University of New Castle.

Werbos P. (1988), *Backpropagation: past and future*, IEEE International Conference on Neural Networks, 343–353, DOI: 10.1109/ICNN.1988.23866.

Wilcoxon F. (1945), Individual comparisons by ranking methods, *Biometrics*, 1, 80–83, DOI: 10.2307/3001968.

Xiaoqiang W. (2022), *Research on enterprise financial performance evaluation method based on data mining*, 2022 IEEE 2nd International Conference on Electronic Technology, Communication and Information (ICETCI), DOI: 10.1109/icetci55101.2022.9832404.

Appendix

The list of selected variables for XGBoost:

1. Changes in provisions per share lagged by 4 quarters
2. Cash and cash equivalents, end of period per share lagged by 4 quarters
3. Operating profit per share lagged by 8 quarters
4. Cash and marketable securities per share lagged by 8 quarters
5. Inventories per share lagged by 4 quarters
6. Long-term receivables per share lagged by 4 quarters
7. Operating expenses per share lagged by 4 quarters
8. Operating expenses per share lagged by 5 quarters
9. Short-term accrued expenses per share lagged by 8 quarters
10. Labour productivity per hour lagged by 4 quarters
11. Retail trade volume index seasonally adjusted lagged by 4 quarters
12. Balance of payment: financial account – net portfolio investments in debit instruments lagged by 8 quarters
13. Central bank: other foreign liabilities: liabilities to non-residents lagged by 8 quarters
14. Residential real estate loans to total loans lagged by 8 quarters
15. Manufacturing selling prices future tendency lagged by 8 quarters
16. Gross margin index lagged by 8 quarters
17. Book value per share lagged by 4 quarters
18. Material fixed asset turnover – trailing 12 months lagged by 5 quarters
19. Net debt / EBITDA (Including Other Operation Income/Expense) – trailing 12 months lagged by 4 quarters
20. Net debt / EBITDA – trailing 12 months lagged by 5 quarters
21. Capex / net sales – trailing 12 months (%) lagged by 4 quarters
22. Net margin – trailing 12 months lagged by 4 quarters
23. Receivables turnover – trailing 12 months lagged by 8 quarters
24. Cash flow return on investment – trailing 12 months lagged by 5 quarters
25. Working capital (net) / sales – trailing 12 months (%) lagged by 5 quarters
26. Change of EBITDA lagged by 4 quarters
27. Change of EBITDA lagged by 8 quarters
28. Current ratio lagged by 4 quarters
29. Days sales in receivables index lagged by 5 quarters
30. EBITDA ROA – trailing 12 months lagged by 4 quarters
31. Equity/assets lagged by 4 quarters.

Figure 1
The example of 34 × 47 picture representing KGHM company

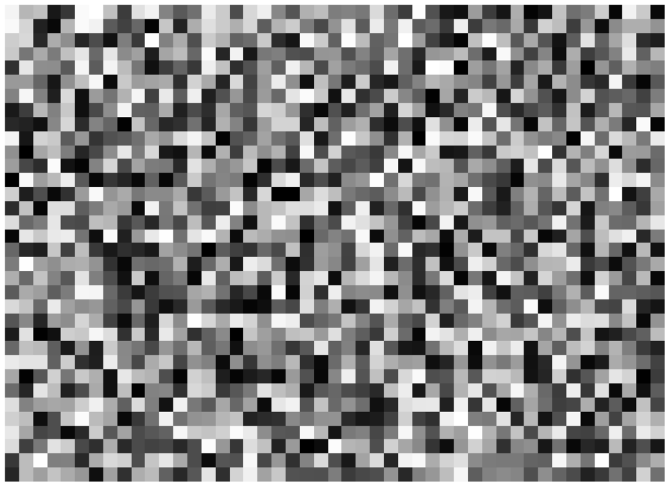

Table 1
Summary statistics on forecast errors for 2019 quarters

Model	Q1 MAAPE	Q2 MAAPE	Q3 MAAPE	Q4 MAAPE	Total MAAPE
SRW	0.658	0.702	0.653	0.736	0.687
XGB_ALL	0.759	0.743	0.779	0.785	0.766
MLP_ALL	1.274	1.290	1.263	1.214	1.260
CNN_ALL	0.794	0.773	0.787	0.786	0.785

Table 2
P-values of the Wilcoxon test of forecast errors for SRW and respective models in 2019

Quarter	Model	XGB_ALL	MLP_ALL	CNN_ALL
1	SRW	0.000	0.000	0.000
2	SRW	0.061	0.000	0.004
3	SRW	0.000	0.000	0.000
4	SRW	0.041	0.000	0.046
All	SRW	0.000	0.000	0.000

Table 3
Summary statistics on forecast errors for all quarters 2017–2019

Model	MAAPE		
	2017	2018	2019
SRW	0.686	0.711	0.687
XGB_ALL	0.821	0.791	0.766
MLP_ALL	0.881	0.786	1.260
CNN_ALL	0.795	0.782	0.785

Table 4
P-values of paired Wilcoxon test of forecast errors for all quarters 2017–2019 and SRW model

Year	Model	XGB_ALL	MLP_ALL	CNN_ALL
2017	SRW	0.000	0.000	0.000
2018	SRW	0.000	0.000	0.000
2019	SRW	0.000	0.000	0.000

Table 5
Summary statistics on forecast errors for RMSE and MAPE in 2019

	SRW	XGB_ALL	MLP_ALL	CNN_ALL
RMSE	0.937	1.191	2.070	1.334
MAPE	0.705	0.958	1.864	1.105

Table 6
P-values of paired Wilcoxon test of forecast errors for RMSE and MAE in 2019

Measure	Model	XGB_ALL	MLP_ALL	CNN_ALL
RMSE	SRW	0.045	0.000	0.000
MAE	SRW	0.024	0.000	0.000

Czy uwzględnienie szerokiego zestawu zmiennych objaśniających jest istotne w prognozowaniu EPS? Dowody z Polski

Streszczenie

W niniejszym artykule zbadano rolę dokładnych prognoz zysków spółek notowanych na giełdzie – kluczową w osiąganiu sukcesu inwestycyjnego, szczególnie na rynkach o ograniczonym pokryciu prognozami przez analityków, takich jak rynki wschodzące, m.in. w Polsce. Podczas gdy analitycy finansowi szeroko prognozują wyniki finansowe spółek na rozwiniętych rynkach, takich jak USA, jedynie niewielka część firm (około 20%) cieszy się podobnym zainteresowaniem w Polsce. Istnieje obszerna literatura poświęcona prognozowaniu zysków na akcję, głównie w USA, choć wyniki tych badań są zróżnicowane. W artykule oceniono dokładność prognoz generowanych przez szeroką gamę zmiennych objaśniających, w tym zmienne księgowe, rynkowe i makroekonomiczne, wykorzystując uczenie maszynowe oparte na drzewach decyzyjnych ze wzmocnieniem gradientowym, wielowarstwowych sieciach perceptronowych i sieciach konwolucyjnych, w porównaniu z sezonowym modelem spaceru losowego. Modele te zastosowano w odniesieniu do danych EPS spółek notowanych na Giełdzie Papierów Wartościowych w latach 2008–2019. W zastosowanych metodach wielowymiarowych wykorzystano kompleksowy zestaw 1598 zmiennych objaśniających, obejmujących specyficzne dla spółki wskaźniki finansowe i rynkowe wraz ze wskaźnikami makroekonomicznymi. Model sezonowego błądzenia losowego wykazał najniższy błąd mierzony za pomocą średniego bezwzględnego błędu arcus tangensa (MAAPE), czego wyniki potwierdzono testami statystycznymi. Liczne kontrole stabilności wyników, obejmujące różne ramy czasowe i wskaźniki błędów, potwierdzają ten wynik. Dominacja modelu uproszczonego może wynikać z tendencji do nadmiernego dopasowania modeli złożonych oraz stosunkowo prostej dynamiki obserwowanej w polskich spółkach giełdowych.

Słowa kluczowe: zysk na akcję, błądzenie losowe, drzewo decyzyjne ze wzmacnianiem gradientowym, wielowarstwowa sieć perceptronów, sieć konwolucyjna, Giełda Papierów Wartościowych w Warszawie