Научная статья

Original article

EVALUATION OF STOCK MARKET EFFICIENCY WITH DEEP NEURAL NETWORKS OF GMDH-type

ОЦЕНКА ЭФФЕКТИВНОСТИ ФОНДОВОГО РЫНКА С ПОМОЩЬЮ ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ ТИПА GMDH

Селицкий Стас, аспирант, Школа компьютерных наук и технологий, Университет Бедфордшира Лутон, LU1 3JU, Великобритания, selitsky@yahoo.com

Дариуш Зегар, аспирант, Школа компьютерных наук и технологий, Университет Бедфордшира Лутон, LU1 3JU, Великобритания

Тусков Андрей Анатольевич, кандидат экономических наук, доцент, заместитель директора по научной работе, Пензенский казачий институт технологий (филиал) ФГБОУ ВО «МГУТУ имени К.Г. Разумовского (ПКУ)», ФГБОУ ВО «Пензенский государственный университет», tuskov@mail.ru

Щанина Елизавета Вячеславовна, студент кафедры «Цифровая экономика», ФГБОУ ВО «Пензенский государственный университет», shchanina03@list.ru

Selitsky Stas, graduate student, School of Computer Science and Technology, University of Bedfordshire Luton, LU1 3JU, UK, selitsky@yahoo.com

Dariusz Zegar, graduate student, School of Computer Science and Technology, University of Bedfordshire Luton, LU1 3JU, UK

Tuskov Andrey, Candidate of Economic Sciences, Associate Professor, Deputy Director for Research, Penza Cossack Institute of Technologies (Branch) of the

Federal State Budgetary Educational Institution of Higher Education «MSUTU named after K.G. Razumovsky (PKU)», Penza State University, tuskov@mail.ru

Shchanina Elizaveta, student of the Department of Digital Economy, Penza State University, shchanina03@list.ru

Аннотация. В исследовании проверяется гипотеза рынка эффективности на двух финансовых рынках, представленных польской и болгарской фондовой биржей. Основной упор делается на использование методов машинного обучения, таких как «неглубокая» прямая связь и «глубокие» нейронные сети с целью проверки слабой формы эффективности фондового рынка. Данные, использованные в этих исследованиях, были собраны из индексов SOFIX и WIG. Исследование польского фондового рынка было сосредоточено на 2007-2009 годах, в то время как изучение болгарского фондового рынка было на данных с 2007 по 2012 год. Результаты показывают явное преимущество нейронной сети глубокого обучения перед авторегрессионными и прямоточными нейросетевыми моделями.

Abstract. We analyse the Efficiency Market Hypothesis on two financial markets represented by the Polish and Bulgarian stock exchange. The main focus is on the use of machine learning methods, such as "shallow" feedforward and "Deep" neural networks in order to verify a weak form of the stock market efficiency. The data used in these studies were collected from the SOFIX and WIG indexes. The study of the Polish stock market focused on the years 2007-2009, while the study of the Bulgarian stock market was on the data from 2007 to 2012. The results show a clear advantage of a deep learning neural network over the autoregressive and feedforward neural-network models.

Ключевые слова: Гипотеза эффективного рынка, машинное обучение, искусственная нейронная сеть, Варшавская фондовая биржа, рыночные аномалии

Keywords: Efficient Market Hypothesis, machine learning, Artificial Neural Network, Warsaw Stock Exchange, market anomalies

Introduction

The Efficient Market Hypothesis (EMH) is considered as a controversial theory since its first official publication in 1964. There are still many opponents of this thesis, however it is a commonly known tool used in order to define the financial markets efficiency. The theory seems to be even more interesting in the case of a financial crisis. Also, the market anomalies are a very intriguing phenomenon which may affect more or less the efficiency of the capital market. There are many methods that can be used in order to verify the EMH, however none of them is stated to be the most suitable and reliable. The interesting approach can be application of machine learning methods to develop forecasting models, which can be useful especially in testing the weak form of efficiency.

The main objective of this paper is to analyse and compare performance of machine learning methods applied in researches aiming to verify the hypothesis of weak form of efficiency of financial markets in years of the world's financial crisis. There were two studies chosen, in which the autoregressive model, feedforward neural network and GMDH-type neural network were used. The first chosen research has examined the weak form of efficiency of the Warsaw Stock Exchange in the period 2007-2009 with the use of machine learning tools in modelling economic models.

The analysis will be based on results achieved by forecasting models, as well as on the research conducted by Paweł Strawiński and Robert Ślepaczuk in 2008. The authors of this work aimed to verify the EMH with the use of high frequency data. This seems to be an interesting approach that can be inspiring for future research.

This paper includes three chapters. In the first one the theoretical background of Efficient Market Hypothesis is presented. There is historical background of the theory, as well as its critics and verifying methods are discussed. Also, the examples of market anomalies are listed and described. Finally, the studies verifying the efficiency of the Warsaw Stock Exchange are shown. The second chapter covers the machine learning methods. Firstly, some issues connected with testing the EMH are presented. Secondly, the Artificial Neural Networks are described, including the discussion about their benefits and

drawbacks. Then, the Group Method of Data Handling is shown with special emphasis on its neural type – GMDH-type neural network. Lastly, the studies in which the machine learning methods were applied are introduced. In the last chapter, the comparative analysis is conducted. In the first section the chosen researches are described. Then, the comparison of performance of these models in both researches is done. Finally, the analysis of the assumptions about efficiency of WSE is undertaken.

In spite of the fact that the theory is more complicated than the EHM, it is not so easy to check the hypothesis with the use of empirical experiments. There are at least 5 factors which have to be specified by the forecasting experiment [12].

The first one is a set of forecasting models, which are accessible at any time, and include methods of estimation. Secondly, forecasting models are selected by the search technology, but only the best or mixture of best can be picked. Thirdly, existing "real time" data sets, including public and private information, and preferably the cost of obtaining this information. The fourth factor is that economic models for the possible danger of premium reflecting economic agent compromise must be in the middle of current and future payoff. The last fifth are available technologies of trading, size of deal cost and limitations on items of the discussed assets.

There are some studies and researches from the past, which are in favor of EMH over random walk, martingale, white noise and fair game. The lack of opportunities of arbitrage shows that investors are able to make extraordinary profits only when miss value stock occurs on a rational stock market. In fact, when it takes place the stock prices are automatically adjusted. The market anomalies, which will be presented later, are one of the reasons why that happens.

In this project multi-approach techniques are going to be used, as they increase the results comparability and the reliability of the project, more than the individual techniques. There are two wide groups in the EMH testing methods: Non-Parametric (e.g. Autocorrelation test) and Parametric tests. When a test of a weak form of efficiency is being done, only past data can be used to forecast the

future prices of stocks. Therefore, testing the hypothesis can be done with the use of empirical analysis. The different ways to accomplish it are:

Firstly, the primary method of empirical evaluation is Random-Walk testing, which follows the behaviour of stock prices. It can be completed with the use of: variance ratios, auto-regression tests, test of normality, test of chaos theory or test of time series. Lastly, the prices follow the random walk, therefore the trends and patterns prediction is impossible.

Second method is the autocorrelation testing, which can be positive or negative. So when positive the EHM is rejected, and vice versa when negative then EHM is verified. The EHM states that between the price movement over the time there should not be any correlation. So it checks if the fluctuations of the current prices are connected with the fluctuations of the past prices.

Run test is another method of testing which allows to overcome some weaknesses of autocorrelation test, as it is able to detect the serial randomness or dependence of the price fluctuations. It looks for movements of prices in a particular time frame, and then compares the forecasted results of random series with the changes that actually took place [10].

The last method of testing is a combination of auto-regression and moving averages, which was described by Box and Jenkins in 1978. The aim of this method is to show that the return rates depend not only on the past prices, but also on the errors in forecasts about the past and on current prices.

The fluctuations on the stock market play a huge role for investors, and that makes the issue very common. There are a number of studies analysing the market anomalies and their possible consequences. There are four main groups that contain the most important anomalies that occur on the capital market: the calendar effects which allow to earn above average return rate in the particular time frame; anomalies linked with the financial rates of companies; effects connected with publishing the important information about the companies; anomalies linked with occurrence of significant correlation coefficients (positive and negative) in different time frames (short, medium and long period).

The examples of the calendar effects are: the month-of-the year effect (here the most popular is the January effect - as it was shown in experiments the return rates in January are statistically higher than in other months), week-of-the year effect, day-of-the week effect (here comes the results of French's research, who found out that the average return rate on Monday is negative, while on other week days it is positive), hour-of-the day effect.

These are the most common anomalies that occur on markets all over the world, however it was proven that they act only temporarily and disappear when some fluctuations on the market appear. That means the anomalies do not cause the ineffectiveness of the market .

Another group of anomalies are the effects linked with financial rates of companies, and these are: P/E (price to earnings ratio) effect (it was proven that stock with low P/E index earn above average return rates in comparison with the average of all tested companies [3], BV/P (book value to price ratio) effect, small-firm effect.

There were also some effects connected with publishing particular information about the companies examined. These information are: finance results, finance results' forecasts, dividend payment or changes in politics of dividend, stock's public offer, split of stock and acquisition of own stock. In the first case, the reactions of companies on finance results were not so clear. The underestimation, as well as overestimation of importance of information about the results were observed. It was shown that the forecasts underestimate the importance of negative data, and overestimate the importance of the positive one, especially in the short period of time, which finally reflects in the valuation of stock in the long period [1]. The same was observed in the case of finance results' forecasts, where negative data was underestimated and the positive data contrary. It resulted in the inappropriate reaction of return rates of particular companies [2]. When it comes to the politics of dividend, it was observed that companies informing about beginning or increasing the amount of paid dividend achieved above average return rates in the following years in comparison to companies informing about ending or decreasing the dividend [13].

These are only a few anomalies existing on stock markets, showing that return rates do not have random distribution which seems to deny the EMH. Obviously, the anomalies often are just temporary and can lead to above average profits, however it still seems quite interesting what may cause the anomalies. Here are the most commonly presented explanations: overreaction of market participants toward the received information, underestimation of importance of received information, herding behaviour. The last one is strictly connected to the overreaction of market participants, as they react together toward received information, no matter what their individual opinion on the information is. Then the received information is repeatedly reflected in prices due to duplicate behaviour of market's participants.

There are also some theories explaining the time anomalies. There is no point in presenting them all, so I have decided to present three most popular hypotheses. First of them is the tax-loss selling hypothesis, referring to U.S. tax law, specifically to the ability of deducting from the capital income from profitable stocks the losses incurred on other stocks. At the end of the year many unprofitable stocks are being sold, so in December the prices dramatically decrease. Whereas in January the stocks become massively bought-back, which causes growth fluctuations and returns to record before declines. The second explanation is so-called *window dressing*. At the end of the year, managers managing investment portfolios, in order to present the best results, sell losses incurred assets. This causes a decrease in stock prices and the most underestimated items are opened again at the beginning of new year. It creates demand which causes an increase of the assets' value. The last hypothesis is the way of rewarding the managers. It is assumed that the wages of managers managing the assets depend on the results of their investments. The stock market index is most commonly the reference measure, however generally their wages increase less than proportionally relative to profits made by them. After reaching a particular point, managers make their profits, because the decline of stocks' prices could expose them to proportionally higher losses in their wages. The settlements are generally regulated at the end of the year, so in order to present big profits in the following period, at the beginning

of new year managers invest in small and risky companies, and that causes demand pressure and increase of their prices [8].

The same period was examined by Jajuga and Papla in 2000, with the use of a few methods, such as random walk tests, randomness tests based on a two-variance test, or Alexander's filters. The authors checked the daily, weekly and monthly return rates. The results were contradictory, however in major part the hypothesis about the weak form of efficiency could have not been rejected. The lack of the day-of-the week effect, as well as the January effect was stated in a statistically significant way.

In 2003 A. Szyszka did research on the years 1991-1999 on WSE. The verification of the weak form of efficiency gave two results. In 1991-1994 the stock prices' fluctuations did not follow the random walk model, and the size of the differences could provide the ineffectiveness of the market. After 1994 the effectiveness was improving, which made the EMH impossible to be denied. There were some small anomalies, similar to these appearing on more mature markets. The outcomes of the analysis of time distribution of return rates were quite intriguing. The return rates for every index used in the study and 29 companies were substantially higher on Mondays than on the other days of week. It was completely different than on the American stock market, where the increases were the smallest on Mondays, and sometimes there were even decreases observed. Another observation done by A.Szyszka was the regularity of negative return rates of the majority of examined companies on Tuesdays. No interesting observations were done on other days of week.

S. Buczek (2005) made an analysis of earlier papers about the Polish stock market, written by J. Czekaj, M. Woś, J. Żarnowski and A. Szyszka, and expanded them with studies on years 2001-2004. Buczek concluded that the assumptions of theory about the weak form of efficiency are too idealistic, as the assets' prices depend on many important interfering factors. Investors can affect the market assets' prices by receiving biased information. This can disturb the effective performance of the market, leading to various anomalies, especially time anomalies. Even though the author claims that the WSE characterizes a high level

of efficiency in the weak form. The verification of the semi-strong form of efficiency did not provide consistent results. According to Buczek, testing this form of efficiency should be followed by considering the size of the company. This assumption was made due to the fact that the behaviour of small companies especially resulted in many deviations, which could provide abnormal profits. Also, the ability of gaining higher than average returns thanks to following press information was observed. These observations led to the assumption that the conditions of semi-strong effectiveness were met only partially. When it comes to the strong form of efficiency, Buczek claims that practically examining this form is impossible, as every investor should behave rationally, and use of confidential data would be illegal.

Machine learning (ML) is one of the most popular fields of Artificial Intelligence (AI) in recent years. It is well-known and commonly used in many fields of science, such as engineering, medicine or economics. The main aim of ML is to construct mathematical models or computers which are able to learn from input data. In economics designing prediction models and testing their performance is done with the use of ML. These models are created to be able to adapt to a rapidly changing environment, which undoubtedly the life stock market environment is. Now most of them are focused on a single learning algorithm which performs in a controlled changing environment. The goal for the future is to create algorithms which will be able to make their own decisions suited to the actual situation.

The "learning" term in the case of intelligent systems differs from human learning. Such systems use some fragmentary knowledge to achieve certain results. What is common to all system is the empirical data entered into the system, on which the learning set is based. It is called an input of the system.

Basically, the machine learning methods are classified by the way of representing knowledge acquired by the system which is learning. Thus, there are two main groups: knowledge oriented methods and black-box methods.

The representation of knowledge that is explicitly created and used in the learning process by the first group of methods, is in the form of symbols, which are

understandable for humans and can be appropriately interpreted by them. The examples of such methods can be grouping algorithms or decision trees. In the second case people are unable to directly interpret the internal system records, as they are usually sets of numbers which reflect distance or weights measures or coefficients. This is why such systems are called the "black box". Artificial Neural Networks are an example of such methods.

There is also a division of machine learning methods into supervised learning and unsupervised learning. In the first case, the system receives a set of examples in the form of two-element vectors. If we label the vector as $\langle x, z \rangle$, where the first variable (x) represents the input information, and the second one (z) represents the sought output information. Thanks to the learning process, the system is able to find the answers 'z' for 'x' data, and thus the representation of knowledge for new input data is created. In the unsupervised learning, the system received only the input vector, and basing only on examples, aims to find regularity in the dataset.

An important feature of machine learning methods is using them in order to construct the classification systems. After the analysis of a set of learning examples and finding a particular unknown function which describes the problem, the classification of new examples to already defined groups occurs. The classifier should be then verified, which can be done with test examples, which were not used earlier. Subsequently, the verification tools are used to evaluate the size of the classification error.

Another way to verify a classifier is to make an experiment, in which there would be division of samples into teaching and test examples. There are different techniques used, depending on the number of cases, and these can be: f-fold cross validation or leave-one-out. In the analysed papers the f-fold cross validation has been used, and the sample was divided into k parts which were equal. Each of the parts was used to evaluate the classifier's accuracy. Then, the sum of results of each iteration was averaged, resulting in one rating of the classifier model [8].

Method

In the following sections the techniques of machine learning are described, and they are: Artificial Neural Networks and the GMDH method, with its type – GMDH-type neural network. These techniques are presented, as they were used in analysed papers in order to verify the efficiency of Polish and Bulgarian stock markets.

Testing EMH problems

The EMH has various advantages and disadvantages, like the other hypothesis, however it is one of the best choices when it comes to making decisions about the stock market. A lot of different studies are done in order to investigate and observe the market efficiency of developing financial markets. All of these studies are showing varied outcomes about the efficiency of the market, and many cannot reach a final assumption about the reliability of EMH. It is due to that specific factors need to appear so the tests could be reliable, and these are: thin trade, nonlinearity of price of assets and the impact of financial liberalization on emerging market performance [10].

Important to mention is that the EMH is at some point dependent on the cost of transaction and restrictions of trading on the markets. Secondly, the most crucial issue here is the cost of transaction. If it is high, and mainly this is the case, then prediction models are useless, no matter if they create proper predictions. Therefore, when the cost of transaction in a given time frame is exceeded, only then EHM can be invalidated by the predicting models.

In the case of choosing the predicting model, it is important to consider the costs of research and transaction. EMH claims that all prediction techniques are ineffective, as even if they are able to forecast the fluctuations of prices, the price of research exceeds the profit. So again it is shown that the market cannot be outperformed by anyone. The predicting models could deny the EMH only when they result in higher returns than the costs incurred by the transaction.

The accuracy is another issue which has to be faced while choosing proper technique for predicting time series. It is crucial when selecting a forecasting model, as it has impact on such issues as lack of information or fitting the datasets pattern. The choice of the most accurate model will result in more precise results of

prediction, and this would make the decision making process easier. The comparison of the prediction results with the actual ones can be a way to check the accuracy. Unfortunately, it is really complicated, due to the need to choose an approach in comparison of the accuracy of various methods. As it was stated by a number of researchers, the out of sample approach is the most reliable while checking the model's accuracy.

About the "random walk" problem we can read about the work of D. Gruen, M. Beechey and J. Vickery from 2000. It is written that the Efficiency Market Hypothesis claims that the financial markets' prices differ randomly relative to the new information. It is impossible to predict their movement and also the additional risk can occur, so no one can be able to outperform the market. Sometimes trends occur, however the individual assets' prices are considered to be moving randomly. Logically it can be assumed that the market prices confirm the EMH theory and they perform accordingly to the "random walk" model. When the particular market is considered to be efficient, it can be expected that the "random walk" model will take place and the movement of prices will only partly reflect the past data.

There are some critics of the theory who claim that there is a short-term sequential correlation and that the occurrence of many successive series, which forms a trend, can reject the occurrence of the "random walk" model on the efficient market [14]. Even though the results of the research are possible, it is not possible to deny the EMH.

Generalization signifies the quality of performance of the network on data samples, i.e. how well the built model fits to the existing set of data. In order to define a network with optimum complexity and reach precise results, this method minimizes chances for error and provides more accurate predicting results. Firstly, to achieve this, target data should be genuinely represented by the training one. Secondly, the NN should have optimum complexity and size considered in the set of data. The reason for that is memorising the information, this means having a problem when dealing with new data, or as we can find in Haykin's paper from 1994, less ability to generalize between input and output. This can lead to huge problems when EMH is tested. In simple words this may happen when the correct

values are not produced after the training due to generalization. Haykin also discusses factors that can influence the generalization, and there are: the size of the training sets, the efficiency of the sets, physical complexity of the problem or the NN design. All of these issues are severely related, particularly the training set of data and the NN model. Generally the number of weights is equal to the number of training sets. However the number of the weights can be increased by the number of delays of that system. Logically, to avoid generalization the weights number should not be larger than the tolerable error on the test multiplied the number of training samples. Different ways and more methods to deal with generalization will be described in future work.

When it comes to the informational adaption of the market, E. Fama in 1969 showed that in an efficient market the prices are immediately adjusted to new data. Thanks to that the EMH gained many followers. Actually, it is not always true, as it has been proven that some data has a delay and prices do not immediately reflect it. This data can be for example financial reports of companies or other data about the capital market. There is research done on this issue, especially focusing on anomalies on the market, as they are the key aspects influencing the adoption of information. Most of the critics of the theory use it in order to show the inefficiency of information, as when there is delay in the reflection of information on the prices, some investors can use it and make unusual profits.

Analysis of investment funds can provide a lot of intriguing information, as they employ managers that spend lots of time in order to collect information. Theoretically, such kind of information should give them a huge advantage in comparison to other players on the market. However, the efficiency of their profits do not differ much from profit made by other players, who use passive strategy. The hypothesis of strong efficiency of market claims, that the return rates of funds being actively managed, will be equal to profits made with the use of passive strategy, before managers' fee. On the other hand, the hypothesis of weak market efficiency says that they will be equal, but after the manager's fee.

It was observed that in the 50s and 60s of the 20th century in the USA the profits made by passive investors were even higher than those from investment

funds. It changed a little bit in the 80s. However, it was a general observation. When it came to individual cases it was noticed that some of the investment funds really make abnormal profits. Mostly, it was funds managed by managers who graduated from very good universities. Investment funds still act in accordance with the EMH, as expert knowledge, as well as professional management incur costs, that offset the possible to achieve abnormal profits [8].

Artificial Neural Networks are computable models constructed in a pattern of neural structure of the brain. They have been constructed in order to tie the abilities of computers, which are irreplaceable in many situations, e.g. performing complex maths, as well as the human brain, whose structure is an extraordinary compound, with its connected to each other over ten billion neurons. The ANN are profoundly explored and applied in many fields, such as biometrics, engineering, optimization, and finally analysing financial time series. Many researchers have been using ANN in order to forecast fluctuations on the stock market. Kimoto with his team was one of the first who applied an Artificial Neural Network to forecast the stock market in Tokyo [7].

Artificial Neural Networks can have a combination of structures, which depends on a specific issue that needs to be solved by them. Usually their distinctions are between the neuron's connections and the cycles between them.

Group Method of Data Handling

The main issue that can appear in the modelling process of complex prediction models in the fields of economics, social sciences, ecology etc. is the bias of the researcher, who moves it onto the model. Many of the studies' results done in those fields are non-consistent and unclear, due to the fact that since the prediction models become so complex, the preliminary assumptions of researchers can be just inaccurate surmises. It was the reason for inventing the GMDH (Group Method of Data Handling) by A. G. Ivakhnenko in 1966. This method enables researchers to construct complex models without the necessity of making any presumptions about their internal architecture. This method makes it possible to build a model that is optimally complex, basing only on data, not the biased assumptions of the researcher. Thanks only to the simple relation of input-output

systems, the algorithm can create a self-organised model that can be applied in many complex systems problems, such as forecasting, synthesis control or identification [17].

The rejection of the deductive approach was not the only assumption of the first version of the model. The second one was the application of the polynomial in the process of creating the structure by variation of partial models. The polynomial degree that results in each iteration is doubled, in accordance with polynomial functions from the earlier step. The least squares method is used then to calculate the optimum values of parameters. This allows to reach the model that would be optimally complex and have complicated structure, just in a few steps.

What differs the GMDH from the regression method is finding the optimum organisation of the structure and applying the internal as well as external criteria of sorting. In the process of creating the model's structure with use of GMDH, the layers are built in sequence – the second layer is built only after the first becomes trained. There are some external criteria, according to which each neuron (consisting of only 2 inputs) is being trained, and then the choice of the best working neurons occurs.

Predicting financial time series

In this section there are presented some studies showing the machine learning models application in predicting financial time series. There are many researches connected with that issue, so here are only a few chosen. Mostly, there are studies in which Artificial Neural Networks are used, as well as the GMDH method.

The first chosen study was conducted by R. Domaradzki who aimed to define investing strategies and predict short-term trends on the Warsaw Stock Exchange in the period 1997-2003, by using 59 artificial neural networks. There were chosen 15 types of variables, which were statistically significant, to model WIG20 index, futures contracts on WIG20 and the KGHM company. In order to predict trends there were three different neural networks used (perceptron, linear and radial). The worst results were made by the linear network, which could have been caused by the fact that the relations on financial markets are non-linear. The

perceptron network was profitable, however it also incurred losses, while the radial network allowed it to gain decent profits (Domaradzki, n.d.).

In the work of A. G. Ivakhnenko and J. A. Muller (1997) describes the idea of the GMDH algorithm, and the possibility of use the GMDH-type neural network on financial markets is presented. The results of modelling the New York stock market for the period February-June 1995. There were seven variables used and the maximum number of delays was set to 35. The results of that study has shown that the GMDH method can be successfully applied in predicting financial time series.

In 2008 in a research conducted by M. Abbod and K. Deshpande the optimized method of GMDH was used to forecast the dollars to euro's exchange rate (USD/EUR) in the period 2004-2007. The method was optimized with use of Genetic Algorithm and Practical Swarm Optimization. First 1000 observations was used to train the algorithms, the following 102 to test the results. In order to calculate the results of models the MAPE and RMSE measures were applied. Even the standard GMDH method gave better results than the normal linear regression. Thus, the additional optimization significantly improved the forecasting results.

In the work of S. H Chen and C. H. Yen (1996) the evolutionary algorithms were applied in order to examine the informational effectiveness. The evolutionary algorithms are one of the methods of machine learning's methods. The authors examined the return rates from Taiwan TAIEX and American S&P500, but from whole observations from 1974-1997 there was only sample of 50 for each of the examined indexes. The data was then divided into training and testing periods (=5), and the MAPE error was calculated by the evolutionary algorithm and then compared to the error of the random walk model. The achieved results were better than the results of autoregressive and random walk models. However it could have been due to that only one small period was chosen. When it came to the bigger learning population, the overfitting problem occurred [8].

In the last section of my work I will compare papers in which the analysis of effectiveness of Polish and Bulgarian stock markets was conducted. In those papers there were some elements of the methodology described by Chen and Yeh

applied, such using the autoregressive models as reference points for other models, and using the MAPE error in order to examine the effectiveness of markets.

Data

In a study by Ciemny the daily return rates were used from the main WIG index. The return rates were turned into logarithmic return rates. The accuracy of predictions and the usefulness of particular models was stated with the use of mean absolute percentage error (MAPE), which can be presented as follows:

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right|$$

A_t – actual index value

F_t – forecasted index value

N – number of observations

The data are used to conduct the cross validation in order to compare the analysed models. In this method the sample is divided into training and testing set. There were 13 series with both sets analysed, and finally the average MAPE error was calculated for each of the models, so they could be compared.

Not only the machine learning models were examined, but also the weak form of efficiency of WSE. In order to do that, the random walk hypothesis was introduced. According to it, the price fluctuations are random and accidental. This hypothesis is said to be stronger than the hypothesis of weak form of efficiency. Thus, meeting its requirements would prove that the market is efficient, at least in weak form. The value of MAPE error for a random walk model is equal 1.00.

The period examined in this paper covered five years – from 2007 to 2012 – including phases before, during and after the world's financial crisis. The main objective of this work was to verify the weak form of efficiency of Bulgarian stock exchange with the use of ANN. Also, the machine learning models were examined, as tools for predicting financial time series. The similarities of main aims of both researchers, as well as use of the same models, made me choose this work as relevant for comparative analysis. These two papers will be compared referring to the performance of machine learning models used in both of them.

In has beed used the daily return rates of SOFIX index, which were transformed into logarithmic rates, with the use of the same formula, as described earlier. The research covered the period from 3rd January 2007 to 15th November 2012 (1410 daily return rates observations). The first return rate was calculated with observation from 22nd December 2006, because it was the last available data from the market's working day. There were 14 series of observations, each containing training sets with 100 observations, and testing sets with 10 observations. In this research the methodology of Chen and Yeh (1996) was also applied.

The analysed data was standardized in the same way as in Ciemny's work. Also, the MAPE error was used to examine the accuracy of each model. The efficiency of Bulgarian stock market was confronted with a random walk model.

The aim of this work was to examine market inefficiencies of the Warsaw Stock Exchange with the use of high frequency data (5-minute returns). The daily data was used as a benchmark only. The study covered the period 2003-2008 for HF data and 1998-2008 for daily data. The index that was used was WIG_{20} index futures. What made me choose this research was its aim to verify the EMH in weak form on WSE in the period, which at least partially overlap the period examined by Ciemny. Furthermore, the use of high frequency data seemed interesting to me and I wanted to confront this approach with the method chosen by Ciemny.

In this research there were used the daily return rates of WIG_{20}, which is a weighted sum of 20 largest companies of WSE from the 2nd February 1998 to 31st March 2008, the 2,547 trading days, as well as the 5-minute data from the 2nd June 2003 to 31st March 2008, the 92,199 five-minute intervals after the outlier correction (Strawiński and Ślepaczuk, 2008). In this paper the robust statistical methods were used, however there is no point in presenting them in detail, as they will be neither analysed, nor compared. The most significant here is the use of high frequency data in order to verify the EMH on WSE. This approach, which can be a very interesting tool for the future, leads to finding some intraday effects which could have an impact on informational efficiency of the market. The assumptions from this research will be described in the following sections.

Results

The results obtained can be divided into general modeling results and detailed results for individual models. In the case of general results, the sum of the MAPE error values for the individual models for the WIG index was computed, which was reached in total in 13 periods. The MAPE error values were compared with the random walk model for which the MAPE error is 1.0. Obtained cross-validation results indicate that the worst performance of all models was presented by an autoregressive model with a mean MAPE of 1.5. The Feedforward neural network achieves an average oscillating score around a value of 1 that is close to the result of a random walk. In this case, the best performance of a GMDH-type neural network was presented, with an average MAPE of 0.83. In addition to the lowest mean MAPE value, the results obtained by the GMDH-type neural network have the smallest values of standard deviation and variance, which shows that the results are evenly distributed around the mean and its lowest variability. Although the autoregressive model obtained a higher MAPE error value from the FNN neural network, its standard deviation and variance values are significantly lower than those obtained by FNN. These results indicate the greatest unpredictability of the FNN neural network. Referring to cross validation, the results achieved by the GMDH model proved to be the best, due to the lowest values of mean MAPE error, standard deviation and variance.

On the other hand, taking into account the MAPE errors in the 13 series, it can be seen that the FNN neural network had the lowest error value eight times, on the second place was GMDH method that has been the best three times, and the autoregressive model twice. However, as already discussed earlier, the FNN neural network has also shown a high level of variance and standard deviation, which puts it in a not as good position as the GMDH model.

When comparing the autoregressive and GMDH models, we come to similar conclusions again. The GMDH-type neural network recorded nine times better results, while the autoregressive model only four times. Furthermore, GMDH turned out to be a much more predictable method as it achieved better cross validation results for all three parameters. The final conclusion that can be drawn

from general modeling results is that for the analyzed 13 periods, only the GMDH-type neural network has achieved interesting results compared to the MAPE model of random walk. Both the lower MAPE error rate and much higher predictability compared to other used models indicate that the GMDH-type neural network is an effective method for analyzing and predicting financial time series

In the case of detailed model analysis, only two out of three were used, i.e. neural network FNN and GMDH-type neural network. Initially, the feedforward neural network was analyzed. The results that were achieved by it were considered unsatisfactory, despite the fact that for many series the error results were very good. As a criterion for the quality of network training, the MSE squared error and the early stop algorithm were used. In the last (13) series of analyzed period, the lowest level of MSE was reached in the 81st iteration. Network learning was then discontinued, and its continuation could lead to overfitting and deterioration of predictive quality. In the last series the error for the test period was lower than for the training period.

The main disadvantage of this model is its large variability. The worst results were obtained in series 5, 7, and 10, when the MAPE error levels were significantly higher than those in other series (in sequence: 2.51, 2.88, 4.44). The results oscillating around the mean were obtained in series 3 and 9 (successively: 1.11, 0, .74). It is noteworthy that in all these cases the downtrend of the WIG index occurred. In the case of an upward trend or a variable over time, FNN has done much better than in the case of series 1, 2, 6, 8, 11, 12, 13, where the network has reached MAPE below 0.2, which is a very satisfactory result. However, referring to the high variability of this model, the MAPE average error value was 1,0009, which prevented the author from rejecting the hypothesis of weak WSE efficiency during the period considered.

Another model discussed was the GMDH-type neural network. In the analysed study, the GMDH model in the first series learned to predict a testing sample with a MAPE2 error of 0.17, while the MAPE1 error for the training sample was 0.87. The structure of this model consisted of 3 hidden layers, so the best results in the series were obtained after 3 iterations. Each neuron had 2 inputs

with a polynomial function. The combination of all three neurons with an output neuron with the same weights led to better results.

In summary, the model structure for the first series included 10 neurons in the third hidden layer, 4 neurons in the second layer and 3 neurons in the first layer. Taking into account how many times an individual input variable has been used, it can be argued that variable 10 (used 4 times) was the most important for predicting the analyzed sequence. Second place was 1, 2 and 7 (all used 3 times). Variable 3 was used 2 times, variable 6 times, and other input variables were not used.

As mentioned earlier, the results of the GMDH model were the least variable compared to the other models. The same situation occurred with each series. The highest MAPE error values were recorded in series 4, 7, 8 and 10. The best results were obtained in series 1 and 13, where the MAPE error value was less than 0.2. It may seem that the results are not impressive, but the smallest standard deviation and mean MAPE error of 0.8347, which is the lower result than the random walk error, make the GMDH-type neural network recognized by the author as having the best prognostics properties from the analysed models.

In the research conducted by Ciemny the verification of weak form of efficiency of WSE in the period 2007-2009 was done with the use of daily return rates of the main WIG index. Machine learning methods were used, such as GMDH-type neural network and feedforward neural network. Their results were later confronted with the results of a random walk model. The random walk model was applied, as it is stronger than Efficient Market Hypothesis, so meeting its requirements can lead to an assumption that the market has at least a weak form of efficiency. Unfortunately, this approach is not perfect, as not meeting these requirements do not lead to deny of the EMH. Thus, this research did not provide a clear answer to the question about the efficiency of the WSE. Even though the GMDH-type neural network resulted in better average value of MAPE error than the random walk model, it is not enough to deny the EMH.

Discussion and Conclusion

The aim of this paper was to analyse and compare research which focused on the verification of Efficient Market Hypothesis on the Polish and Bulgarian stock markets in the years of the world's financial crisis. The EMH is a well-known method for defining the efficiency of financial markets. There are many critics to the hypothesis and market anomalies that can affect the effectiveness of stock markets. Even though the theory still remains valuable and useful for analysing financial markets.

The first analysed paper focused on the verification of weak form of efficiency of the Warsaw Stock Exchange in 2007-2009. The machine learning methods, such as GMDH-type neural network and feedforward neural network were applied. The first model achieved better results with lower average MAPE error and better values of standard deviation and variance. GMDH-type neural network as the only one achieved lower MAPE error than the random walk model. This result indicates that this method has great prognostic abilities and high predictability. Unfortunately, the results of this research do not answer the question about the efficiency of the Warsaw Stock Exchange. The comparison of predicting model results to the random walk model does not deny the hypothesis, however it is not enough to prove it. This is why another research was chosen, which aimed to verify the efficiency of the WSE in years 2005-2008 with the use of high frequency data. This research revealed the occurrence of some intra-day anomalies, such as hour-of-the-day effects which can strongly affect the effectiveness of the Polish stock market. However, this research also is not enough to state if the WSE reveals at least a weak form of efficiency. The approach of the use of high frequency data connected to the application of machine learning methods can be a valuable inspiration for the future.

The research verifying the efficiency of Bulgarian stock exchange revealed the same patterns. Here also the GMDH-type neural network performed best with the lowest average MAPE error, as well as the values of standard deviation and variance. However, what is interesting, the feedforward neural network performed much better in this study, providing lower standard deviation and variance. It was

concluded that the longer training period could result in increasing the predictability of the FNN model. Even though the GMDH-type neural network's results were satisfactory, they did not allow the hypothesis of a weak form of efficiency of Bulgarian stock market.

To conclude, the results achieved by machine learning models used in analysed research were satisfactory, however not enough to answer the question about the efficiency of verified financial markets. To make it possible, deep research is needed, with the analysis of market anomalies, also the intra-day effects which can affect the efficiency of the market. The approach of using high frequency data connected to the application of machine learning methods, can be a solution.

References

1. Abarbanell, J. and Bernard, V. (1992). Tests of Analysts' Overreaction/Underreaction to Earnings Information as an Explanation for Anomalous Stock Price Behavior. *The Journal of Finance*, 47(3).

2. Ball, R. (1978). Anomalies in relationships between securities' yields and yield-surrogates. *Journal of Financial Economics*, 6(2-3).

3. Basu, S. (1977). Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis. *The Journal of Finance*, 32.

4. Beechey M., Gruen D., Vickery J. (2000). *The efficient market hypothesis: A survey.* Reserve Bank of Australia.

5. Fama E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance.*

6. Haykin S. (1994). *Neural networks: A Comprehensive Foundation.* New York: Macmillan College Publishing Company.

7. Kimoto T. & Asakawa K. (1990). *Stock Market Prediction System with Modular Neural Networks.*

8. Ciemny, M. (2009). *Weryfikacja hipotezy efektywności rynku w okresie kryzysu finansowego 2007-2009 przy użyciu metod uczenia maszynowego.* Warszawa: Uniwersytet Warszawski.

9. French, K. (1980). Stock returns and the weekend effect. *Journal of Financial Economics*, 8.

10. Oprean, C. (2012). Testing the financial market informational efficiency in emerging states. *Review of socio-economic research*, 4.

11. Sardar, M. (2004). Empirical Finance: Modelling and Analysis of Emerging Financial and Stock Markets. *Springer Science & Business Media*.

12. Timmermann, A. and Granger, C. (2004). Efficient market hypothesis and forecasting. *International Journal of Forecasting*, 20.

13. Malkiel, B. (2003). The Efficient Market Hypothesis and Its Critics. *Journal of Economic Perspectives*, 91.

14. Lo, A. (2004). The adaptive market hypothesis: market efficiency from an evolutionary perspective. *Journal of portfolio management*, 30, pp.15-29.

15. Fama, E. (1991). Efficient Market Hypothesis: II. *Journal of Finance*, 46.

16. Mills, T. (2002). *Forecasting financial markets*. Cheltenham: E. Elgar.

17. Farlow, S. (1981). The GMDH Algorithm of Ivakhnenko. *The American Statistician*, 35(4).

18. Ivakhnenko, A. and Mueller, J. (1997). *Recent Developments of Self-Organising Modeling in Prediction and Analysis of Stock Market*.

19. Varahrami, V. (2012). Good Prediction of Gas Price between MLFF and GMDH Neural Network. *International Journal of Finance and Accounting*, 1(3).

20. Jakaite, L., Schetinin, V., Hladuvka, J., Minaev, S., Ambia, A., Krzanowski,W.: Deep learning for early detection of pathological changes in x-ray bone microstructures: case of osteoarthritis. Scientific Reports 11 (2021). 11. Jakaite, L., Schetinin, V., Maple, C.: Bayesian assessment of newborn brain maturity from two-channel sleep electroencephalograms. Computational and Mathematical Methods in Medicine pp. 1–7 (2012).

21. Jakaite, L., Schetinin, V., Maple, C., Schult, J.: Bayesian decision trees for EEG assessment of newborn brain maturity. In: The 10th Annual Workshop on Computational Intelligence UKCI 2010 (2010).

22. Jakaite, L., Schetinin, V., Schult, J.: Feature extraction from electroencephalograms for Bayesian assessment of newborn brain maturity. In:

24th International Symposium on Computer-Based Medical Systems (CBMS). pp. 1–6 (2011). 14. Jeon, S., Hong, B., Chang, V.: Pattern graph tracking-based stock price predictionusing big data. Future Generation Computer Systems 80, 171 – 187 (2018)

23. Nyah, N., Jakaite, L., Schetinin, V., Sant, P., Aggoun, A.: Evolving polynomial neural networks for detecting abnormal patterns. In: 2016 IEEE 8th International Conference on Intelligent Systems (IS). pp. 74–80 (2016).

24. Nyah, N., Jakaite, L., Schetinin, V., Sant, P., Aggoun, A.: Learning polynomialneural networks of a near-optimal connectivity for detecting abnormal patterns in biometric data. In: 2016 SAI Computing Conference (SAI). pp. 409–413 (2016).

25. Schetinin, V., Jakaite, L., Krzanowski, W.: Bayesian averaging over decision tree models: An application for estimating uncertainty in trauma severity scoring. International Journal of Medical Informatics 112, 6 – 14 (2018).

26. Schetinin, V., Jakaite, L., Krzanowski, W.: Bayesian averaging over decision tree models for trauma severity scoring. Artificial Intelligence in Medicine 84, 139–145 (2018). https://doi.org/https://doi.org/10.1016/j.artmed.2017.12.003

27. Schetinin, V., Jakaite, L., Krzanowski, W.: Bayesian learning of models for estimating uncertainty in alert systems: Application to air traffic conflict avoidance. Integrated Computer-Aided Engineering 26, 1–17 (2018).

28. Schetinin, V., Jakaite, L., Nyah, N., Novakovic, D., Krzanowski, W.: Feature extraction with GMDH-type neural networks for EEG-based person identification. International Journal of Neural Systems (2018). https://doi.org/https://doi.org/10.1142/S0129065717500642

29. Schetinin, V., Jakaite, L., Schult, J.: Informativeness of sleep cycle features inbayesian assessment of newborn electroencephalographic maturation. In: 2011 24th International Symposium on Computer-Based Medical Systems (CBMS). pp. 1–6 (2011).

30. Schetinin, V., Jakaite, L.: Extraction of features from sleep EEG for Bayesian assessment of brain development. PLOS ONE 12(3), 1–13 (03 2017).