ORIGINAL PAPER

Contact to corresponding author: nijole.maknickiene@vgtu.lt, Department of Financial Engineering, Faculty of Business Management, Vilnius Gediminas Technical University, Lithuania

Nijole Maknickiene
Vilnius Gediminas Technical University, Lithuania

Indre Lapinskaite
Vilnius Gediminas Technical University, Lithuania

Algirdas Maknickas
Vilnius Gediminas Technical University, Lithuania

Application of ensemble of recurrent neural networks for forecasting of stock market sentiments

Abstract

Research background: Research and measurement of sentiments, and the integration of methods for sentiment analysis in forecasting models or trading strategies for financial markets are gaining increasing attention at present. The theories that claim it is difficult to predict the individual investor's decision also claim that individual investors cause market instability due to their irrationality. The existing instability increases the need for scientific research.

Purpose of the article: This paper is dedicated to establishing a link between the individual investors' behavior, which is expressed as sentiments, and the market dynamic, and is evaluated in the stock market. This article hypothesizes that the dynamics in the market is unequivocally related to the individual investor's sentiments, and that this relationship occurs when the sentiments are expressed strongly and are unlimited.

Methods: The research was carried out invoking the method of Evolino RNN-based prediction model. The data for the research from AAII (American Association of Individual Investors), an investor sentiment survey, were used. Stock indices and sentiments are forecasted separately before being combined as a single composition of distributions.

Findings & Value added: The novelty of this paper is the prediction of sentiments of individual investors using an Evolino RNN-based prediction model. The results of this paper should be seen not only as the prediction of the connection and composition of investors' sentiments and stock indices, but also as the research of the dynamic of individual investors' sentiments and indices.

Introduction

The subject of the behavior of individual investors is relevant and has been analyzed extensively in recent decades. From the standard, classical behavioral models, which state that it is possible to calculate and predict the direction of market changes and movements, and that the investor is a rational individual, new theories that contradict the old views have emerged over time. The theories that claim it is difficult to calculate and predict the market behavior or individual investor's decision making claim that an individual investor is an irrational individual, which causes market instability. Over time, the existing instability increases the need for scientific research. In recent decades, a number of studies were dedicated to not only determining the correct way to calculate and predict the direction of market changes, but also to find the causes of market instability. Recently, the active use of social media in politics for public purposes has been noted. One influential politician's post on Twitter can cause significant financial market fluctuations. Therefore, the research and measurement of sentiments, the selection of indices and the integration of methods for sentiment analysis in forecasting models or trading strategies are becoming increasingly important for financial markets. Even more questions arise regarding the use of computational intelligence for forecasting financial markets. Are neural network-based models able to identify the financial market's fluctuations caused by politicians' tweeting? Could sentiment index data play a supervisory role for neural network forecasting models?

This study strives to investigate the informativeness of the combination of two distributions that differ in nature, namely the forecasting of stock indices and the forecasting of sentiment surveys (AAII). Stock indices and sentiments are forecasted separately before being combined as a single composition of distributions, which can provide additional information for the investor at the time of decision making.

The aim of this paper, which uses an Evolino RNN-based stacking ensemble prediction model, is to establish a link between individual investor's

behavior, which is expressed as sentiment in this paper and the market dynamic, which is evaluated via the stock market. This article hypothesizes that the dynamic in the market is unequivocally related to the individual investor's sentiments, and that this relationship occurs when the sentiments are expressed strongly and exceed bounds. The novelty of this paper is the prediction of the sentiments of individual investors using the Evolino RNN-based prediction model, which had previously only been used in the exchange market.

The data for the research by AAII (American Association of Individual Investors), an investor sentiment survey, were used. The results of the research are presented as two outcomes. The first is the research of dynamic of individual investor's sentiments and indices, while the second is the prediction of the connection between investors' sentiments and stock indices.

The paper is organized as follows: in the next section, scientific literature is reviewed, further the research methodology — a prediction model based on the ensemble of Evolino RNN, — is described. In section 3, sentiments data are presented and explained. Section 4 presents and discusses the main results. The last section is dedicated to conclusions.

Theoretical background

Individuals' investment behavior has been explored via a large body of empirical studies over the past three or four decades (Jagongo & Mutswenje, 2014). One of the most important features of investors' behavior is the ability to make decisions. The different psychological states of investors impact on their decisions in the finance market. Sahi (2012) explained the idea that the standard model of decision making was based on the belief that human beings are rational agents (Markowitz, 1952), while the standard economic theory states that people make decisions that maximize their utility and that "emotions are automatic processes associated with strong positive or negative utility" (Cohen, 2005), thereby endorsing the belief that people act rationally and take long-term views when making decisions. However, this is not borne out when people actually make real-life decisions (Simon, 1957; Tversky & Kahneman, 1974; Kahneman & Tversky, 1979). In Loewenstein *et al.* study (as cited in Lucey & Dowling, 2005), who developed the risk-as-feelings hypothesis, the idea that "every aspect of the decision-making process is influenced by the feelings of the decision-maker" was stated. Sahi (2012) added that risky decision making is influenced by emotions in many different ways and raised the question of

why the actual behavior of people deviates from the axioms of rationality. Some theoretical approaches concerning rationality and stability should be reviewed.

In Haugen's (2001) view, models based on rational economic behavior cannot explain important aspects of market behavior, which provides clear evidence that the stock market is inefficient. However, substantial empirical evidence suggests that the market may not be completely efficient in its processing of public information (Aboody *et al.*, 2002). The evidence presents a challenge to the efficient markets theory because it suggests that, in a variety of markets, sophisticated investors can earn superior returns by taking advantage of underreaction and overreaction without bearing extra risk (Barberis *et al.*, 1998). According to Fama (1970), in general terms, the ideal is a market in which prices provide accurate signals for resource allocation: that is, a market in which firms can make production-investment decisions, and investors can choose among the securities that represent ownership of firms' activities under the assumption that security prices fully reflect all available information at any given time. Sheila (2010) explained market stability and the rise of instability; according to the efficient market hypothesis, markets are normally stable. There is motion as new information emerges, but market behavior ensures the stabilization of markets around equilibrium. However, there is sometimes instability, which arises from some initial imperfection in the market process that becomes amplified.

Some authors (Maghrebi *et al.*, 2014) have linked the market's volatility to financial instability, and have investigated whether changes in implied volatility in international financial markets can be defined as a function of market sentiment and a realignment process following forecast errors consistent with rational expectations. Shu and Chang (2015) analyzed the interrelationship of investors' sentiments and the finance market's volatility. Baker *et al.* (2012) found evidence that investor sentiment plays a significant role in international market volatility and generates return predictability that is consistent with corrections of overreaction. Antweiler and Frank (2004) found that message bullishness and volume help to predict market volatility, but have limited ability to predict returns. The relationship between investors' positive sentiments and the overconfidence effect was investigated by Huang and Goo (2008). They divided sentiments in natural environment happiness and investment environment happiness and found that when natural environment happiness is stronger investors are less likely to have overconfidence. Recent studies are strongly concentrated on the relationship between investment sentiment and stock market returns. Yang *et al.* (2017) and Ryu *et al.* (2017) in their research found out that invest-

ment sentiment has an influence on stock market returns. Smales (2017) analyzed the cross-sectional effect of sentiment on stock returns by considering industry level returns, as well carefully identifying how the sentiment–return relationship is affected by the state of the economy and demonstrated a strong relationship between investor sentiment and stock returns that is consistent with theoretical explanations of sentiment. Anusakumas *et al.* (2017) analyzed the link between stock returns and sentiment more precisely, distinguishing sentiments as stock specific and market wide. Despite some interpretations in results, the authors have confirmed that the effect of sentiment on stock returns persists.

Wurgler (2007) presented the assumption that investors are subject to sentiment. Investor sentiment, broadly defined, is a belief about future cash flows and investment risks that is not justified by the facts at hand. There are various definitions related to sentiment; Baker and Wurgler (2006) explained it as the propensity of investors to speculate, which includes waves of optimism and pessimism. These waves could also be seen as "expectations of market participants relative to a norm: a bullish (bearish) investor expects returns to be above (below) average, whatever average may be" (Brown & Cliff, 2004). Antweiler and Frank (2004) categorized existing market sentiment measures into two popular classifiers — Naive Bayes and Support Vector Machine — which are employed to classify stock messages into three categories, namely bullish, bearish and neutral.

Prediction can be explained by three paradigms. The first is that the future is the consequence of a past event. The second is that the future is a continuation of the past, while the third is that the future is a distribution of possibilities. Gneiting and Katzfuss (2014) formalized and studied probabilistic forecasting. Gneiting and Ranjan (2013) investigated combinations of predictive distributions and their aggregation methods from the perspective of calibration and dispersion. As sentiment is associated with different attributes (Bank & Brustbauer, 2014), the prediction of sentiment has been analyzed in different ways by various authors as a prediction of stock prices and exchange rates by using investors' sentiments (Rechenthin *et al.*, 2013). Heiden *et al.* (2013) discovered that institutional sentiment significantly predicts returns over medium-term horizons in the EUR/USD market, while research by Qian and Rasheed (2007), Kavussanos and Dockery (2001), Butler and Malaikah (1992), and Gallagher and Taylor (2002) showed that stock market prices do not follow a random path, and can be predicted to some degree. Plakandaras *et al.* (2015) investigated investors' sentiments in terms of the source of additional information in the prediction and direction of change in the exchange market.

There are many scientific works dedicated to stock market prediction using investors sentiment, but the short overview of models for forecasting will focus on the last decade and begins with Wang *et al.* (2013), who proposed a novel sentiment-based hybrid model (SLNM), and used daily financial news and support vector regression to train and forecast non-linear patterns of stock return series. The incorporation of investor's sentiments in a finance market's volatility model and the improvement of forecasting accuracy was described in a paper by Seo and Kim (2015). Cosma and Acampora (2016) provided an overview of recent fuzzy-based approaches to the sentiment analysis of customer reviews. The developed model includes methods for preparing the dataset, extracting the best features for prediction, and the customer review rating prediction. Fuzzy systems can be used without a set of historical data, which is very important in sentiment analysis. Wang *et al.* (2016) proposed an unsupervised multi-granularity fuzzy computing model for the classification of the sentiment intensity of reviews. Computational model for stock market in order to analyze market self regulation abilities, market efficiency and determinants of emergent properties of the financial market, was proposed by Rutkauskas and Ramanauskas (2009), but computational intelligence models are also used for sentiment prediction. A text analysis of more than 900 brief news was used for a prediction model based on Thomson-Reuters' neural network developed by Heston and Sinha (2015). The findings were that positive news stories increased returns quickly, but negative stories had a long-delayed reaction, and were important for investors. Ho and Wang (2016) developed an Artificial Neural Network model to predict the stock prices.

Research methodology. Prediction model based on the Evolino RNN ensemble

An Evolino RNN-based prediction model (Figure 1) was developed as support system for investors in the exchange market (Maknickiene & Maknickas, 2013a; Maknickiene & Maknickas, 2013b). The prediction of the sentiments of individual investors is a new area for researches. Data related to sentiments are expressed as $log_2 (Post/Negt)$.

The prediction model has two inputs. The first one uses historical data pertaining to the financial instrument that we want to predict. The second input uses historical data related to gold prices in USD (United States dollars). The data ranges that were selected are closer to orthogonal inputs. This reduced the calculation time and increased the accuracy of the predic-

tion (Maknickas & Maknickiene, 2012). The ensemble is composed of 176 Evolino RNNs predicting simultaneously (Maknickiene & Maknickas, 2013b). This process requires multi-core hardware resources for timely data processing using MPI library-based parallel computation. The result of the prediction was presented as a distribution of the expected values. The multimodal distribution is more informative than is the point prediction. The calculation of probabilities for the growth and decline of financial instruments (exchange rate, index, stock price) assists in portfolio construction.

Stock market prediction method is based on composition of two distributions: stock indexes predictions and sentiment predictions. Both distributions of predictions are obtained from Evolino RNN-based prediction model. For the first input of model the data of stock market indexes are used and for second input the historical data of gold prices are used. The result is the distribution of expected values. For prediction of sentiments of individual investors, the historical data from AAII are used, and the prediction model generates the second distribution. Exactly the composition of these distributions is the basis for a new prediction method, which gives additional information for investor. The visualization of financial market prediction method using sentiment data is presented in Figure 2. Distribution of finance indexes is black, distribution of sentiments is grey, the last known value marked by dotted line. The columns' height of distributions describes certain frequency of predictions whose values are increasing steadily in the horizontal direction. Neutral sentiment value corresponds to the last known value.

The composition of distributions has been evaluated by the last known value (dotted line in Figure 2). If the main part of both distributions is on the right from the last known value (Figure 2 a), the prediction shows growth of the finance index. If the main part of distributions is on the left (Figure 2 b), the prediction shows fall of finance index. In that case, when the last known value is between modes of distributions (Figure 2 c), the prediction of growth and fall has almost equal probabilities and the decision of investment is very risky. The next case (Figure 2 d) explains the role of sentiments in prediction of finance index. When sentiments are highly positive or low negative (more than the distance of standard deviation) the prediction of finance index is growth or fall accordingly. Especially the distribution of sentiments is important when the forecast of financial index is uncertain, multi-distribution is on both sides of the last known value. When prediction of sentiments is expressed strongly and goes out of limits, finance indexes will also react strongly. This forecasting method is easy displayed by Evolino RNN-based prediction model. The decision in finance market is made by quick classification of parts of distributions on positive-

negative (sentiments) and on growth-fall (finance indexes). The investor can integrate this information in trading strategy or investment portfolio construction.

Sentiments of individual investors

In this paper, the data from AAII (American Association of Individual Investors) were used. The AAII is an investor sentiment survey that measures the percentage of individual investors on the stock market for six months in the future and classifying them into categories of bullish (*Post*), bearish (*Negt*) or neutral (*Neutr*). All individuals are polled from the ranks of the AAII membership on a weekly basis (http://www.aaii.com/sentime ntsurvey). Only one vote per member in each weekly voting period is accepted.

The first step was to classify the survey data into neutral and sentimental data. The classification was done in the following way: If *Neutr* > 50 %, the market is neutral; if *(Post + Negt)* > 50%, the market has polarity and is sentimental. The second task was to classify the polarity. Emotion was determined with reference to Xu *et al*. (2012), as *Emot = Post – Negt,* and the ratio of positive sentiment to negative sentiment as *Post/Negt*. The indicator log_2 *(Post/Negt)* reflects the ratio of positive sentiment indicators to negative sentiment indicators in a given time frame t, ranging from 0 to infinite. When log_2 *(Post/Negt)* > 0, the overall public sentiment is positive; the larger the value of the ratio, the higher the intensity of positive sentiments. Conversely, when log_2 *(Post/Negt)* < 0, the overall public sentiment is negative and the smaller value thereof, the higher the intensity of the negative sentiments.

A comparison of the dynamic in the data for the years 2010–2016 related to individual investors' sentiments and the SPX500 index is presented in Figure 3. The sentiment data are expressed on a log_2 *(Post/Negt)* scale.

The zones of high positivity, weak sentiment and low negative survey opinions of individual investors are marked. The zone of weak sentiments is defined by standard deviations on both sides of the neutral line. The high positivity zone is above one standard deviation from the neutral line, while the low negative zone is below one standard deviation from neutral line.

The dynamic of the weekly historical data in the AAII and NAS100 index for the years 2010–2016 are presented in Figure 4.

The NAS100 line showed a strong positive trend in the period from week 1 to week 263 of 2010–2015. The AAII line was mainly between the (-) one standard deviation and the (+) one standard deviation lines, or above

them. The measure of sentiments later fell to below (-) standard deviation and the NAS100 values also fell.

Historical data for AAII and US30 are presented in Figure 5.

The classification of neutral-sensitive and positive-negative opinions in the survey of individual investors allows for the identification of singularities in the dynamics of the SPX500, NAS100 and US30 indices. When sentiments are highly positive, the indices grow; when they are low negative, the indices fall, and when sentiments are weak, all the indices have stable trends. Thus, the changes of opinion in the survey of individual investors reflects the behavior of the stock market.

Investigation of predictions of the stock market based on sentiment survey data

The prediction model based on the Evolino RNN ensemble has been tested in the exchange and stock markets for several years. We investigated the accuracy of different variations of model, different trading strategies and various portfolio construction methods. The behavior of individual investors is new area of research, which heavily reliant on selecting data not only for prediction, but also for supervision. After numerous tests, historical data from the AAII investor sentiment survey were selected for prediction, while the gold price in United States dollars was chosen for supervision. The prediction of individual investors' sentiments and stock indices allows for the investigation of the composition of distributions, the opportunities to predict and the accuracy of the predictions.

Prediction of individual investors' sentiments

The prediction of individual investors' sentiments was performed using an Evolino RNN ensemble. The ratio of bearish and bullish historical data from www.aaii.com was used for the first input. For the second input, the historical data regarding gold prices in USD were used. The selection of inputs was determined by seeking historical data that are close to orthogonal ranges for both inputs. The result of the prediction is the distribution of the expected ratio of investors' sentiments (Figure 6). Weak sentiments are defined by standard deviations on both sides of the neutral line (0).

The high positivity zone is to the right of one standard deviation from the neutral line (0,6), while the low negative zone is to the left of one standard deviation (-0,6). Figure 6 presents the distributions of two weeks:

black distribution at 12–11–2015, grey — 19–11–2015. The composition of the last-known values on 05–11–2015 and 12–11–2015 (black vertical lines), predicted distributions and real value on 19–11–2015 (dotted line) show the correct direction of predictions.

The prediction was presented using a logarithmic scale for improved presentation. The dynamic of investors' sentiments was shown; the value of the ratio decreases, but remains in the optimistic area. The probability that the AAII ratio will grow on 05–11–2015 is 21%; that will fall is 79%. On 12–11–2015, the probabilities are 44% and 56%, respectively. Real values show the correct direction of predictions.

The distribution of the expected values of sentiments provides information about changes in the sentiments of individual investors. It indicates that the three main predicted states of sentiments are neutral, negative and positive. The predicted weak sentiments are determined when the most probable values are in the -0.6 to 0.6 range. Highly negative (<-0,6) and highly positive (>0,6) sentiments can predict the direction of the change of the stock indices.

Combination of the AAII and stock indices predictions

The decision was made to compare the predicted distribution of the expected ratio of investors' sentiments with the predictions of three indices on the financial market, namely the weekly NASDAQ (NAS100USD1008), the weekly Dow Jones 30 (US30USD1008) and the weekly SPX 500 (SP500USD1008).

In Figure 7, the composition of AAII and one NASDAQ distribution of predictions is shown. The last-known value of the index is set equal to the bullish/bearish ratio zero point. This values marks the decision-making moment.

At the moment of prediction, or close of market, investors are in receipt of the last-known value of the index. The prediction of the NASDAQ index shows the growth of the index in the following week, because the main part of the distribution of the expected values of the index (Figure 6, grey) is to the right of the last-known value of the index. The main part of the distribution of the prediction of investors' sentiments is also to the right of 0, which means that it will be positive. The probability that the NASDAQ index will grow is 65%, while the probability that it will fall is 35%. The probability that investors will be optimistic in the following week is 75%, and that they will be pessimistic is 25%. The real value (dotted line in Figure 7) of the index shows the correct prediction.

The combination of the distributions of the Dow Jones 30 index and the AAII sentiment survey (Figure 8) is done in the same way.

The distribution of the prediction of the expected values of the Dow Jones 30 is on both sides last-known value; thus, the probability of expected growth is 50%, and the probability of an expected fall is 50%. However, the prediction of the sentiments of investors is optimistic. The real value (dotted line in Figure 8) shows that the index is growing.

The combination of the distributions of the SPX500 index and AAII (Figure 9) is done in the same way.

The main part of the distribution of the expected values of the SP500 is to the right of the last-known value; thus, the probability of growth is 58%, while the probability of falling is 42%. The optimistic prediction in the AAII survey only reinforces the growth forecast. The real value (dotted line in Figure 9) shows that the index is growing.

Estimating the accuracy of classification

Two reasons for the importance of the estimation of accuracy of the classification model were suggested by Sharda *et al.* (2013). The first is that accuracy can be used to estimate the prediction accuracy. The second is that it can be used to choose the best classification model. Our classification problem involves three decisions: positive, neutral and negative sentiments or buy, keep, sell decisions on the finance market. The accuracy metrics of predictions are expressed through overall classifier accuracy (Sharda *et al.*, 2013):

$$\text{OCA} = \frac{\sum_{i=0}^{n}(True\ Classification)}{Total\ Number\ of\ Cases} \tag{1}$$

where: OCA is overall classification accuracy; n – number of decisions; i – the number assigned to the decision.

A statistical investigation of the accuracy of weekly predictions was conducted for four months; some predictions were not useful because the parameters of distribution were wrong. The comparison of the accuracy of predictions is presented in Table 1.

The improved accuracy (87%) of the sentiment prediction is caused by less volatility in the data pertaining to sentiments in comparison to the volatility of finance indices. The combination of AAII and the stock indices has

a prediction of 57% accuracy when all predictions of sentiments are used, regardless of the sentiments' clarity. The accuracy rose to 73% when the predictions that sentiment is weak were rejected: between (-) one standard deviation and (+) one standard deviation. The forecasting accuracy of the combination of distributions of AAII and the stock indices is less than is the accuracy when considering sentiments and stock indices separately, but an investor has additional trading information and the choice to develop new marketing strategies and models.

Conclusions

The opinions of individual investors expressed via the sentiment survey should be seen as an important indicator reflecting possible future changes in the finance market. The classification of neutral-sensitive and positive-negative opinions of individual investors, which was determined via a survey, shows the singularities in the dynamics of the SPX500, NAS100 and US30 indices. When sentiments are highly positive, the indices grow; when sentiments are strongly negative they fall, and when sentiments are weak, all the indices have stable trends.

The results of the predictions of individual investors' sentiments (AAII), which were made using an ensemble of Evolino RNN, and different stock indices are expressed as probabilities. An informative prediction tool for investors could be the combinations of distributions of individual investors' sentiments and stock indices when the values are indicated at the moment of decision making. When the distributions of investors' sentiments and stock indices are the last-known value or the closing value of the previous day, the decision becomes stronger. When the distributions are on both sides of this value, the decision could be called risky. When the prediction that the sentiment would be weak were rejected, the accuracy rose to 73%, between (-) one standard deviation and (+) one standard deviation from the neutral line.

The results of this research confirm the hypothesis of this paper, namely that dynamics in the market are unequivocally related to individual investors' sentiments and that this relationship occurs when the sentiments are expressed strongly and exceed boundaries.

The new field of the application of the Evolino RNN ensemble and the new application method provide additional information for investors in the stock market.

Although the research is prepared and carried out carefully, some limitations still need to be mentioned. First of all, the research was conducted in

the four different groups, each group lasted for 15 weeks. Fifteen weeks is a minimum time interval for researchers to observe most cases of impact of sentiments to the changes of exchange rates. A longer time period of investigation might bring better results. Secondly, the prediction invoking the Evolino RNN-based prediction model, requires extensive time and calculation resources, so the ability to repeat the same prediction is limited. Thirdly, only one sentiment index, the historical data from AAII, was investigated. Nowadays the number of sentiment indexes arise rapidly and comparison of different sentiments indexes can provide additional information. The upgrade of computational resources, additional investigations of other sentiment indexes and sentiment measures would be very useful. Intended direction of the future research is a construction of investment portfolio invoking the ensemble of Evolino RNN with inclusion of sentiment data.

References

AAII (2017) Investor Sentiment Survey. Retrieved from http://www.aaii.com/sentimentsurvey.

Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *Journal of Finance, 59*(3). doi: 10.1111/j.1540-6261.2004.00662.x.

Anusakumar, S. V., Ali, R., & Hooy, C. W. (2017). The effect of investor sentiment on stock returns: insight from emerging Asian markets. *Asian Academy of Management Journal of Accounting and Finance, 13*(1). doi: 10.21315/aamjaf2017.13.1.7.

Aboody, D., Hughes, J., & Liu, J. (2002). Measuring value relevance in a (possibly) inefficient market. *Journal of Accounting Research*, *40*(4). doi: 10.1111/1475-679X.00078.

Baker, M., & J. Wurgler. (2006). Investor sentiment and the cross-section of stock returns. *Journal of Finance, 61*(4). doi: 10.3386/w10449.

Baker, M., *Wurgler*, J., & Yuan, Y. (2012). Global, local, and contagious investor sentiment. *Journal of Financial Economics, 104*. doi: 10.1016/j.jfineco.2011 .11.002.

Bank, M., & Brustbauer, J. (2014). Investor sentiment in financial markets. Retrieved from http://www.uibk.ac.at/ibf/mitarbeiter/bank/investor-sentiment-in-financial-markets.pdf.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics, 49*. doi: 10.3386/w5926.

Brown, G. W., & Cliff, M. T. (2004). Investor sentiment and the near-term stock market. *Journal of Empirical Finance, 11*. doi: 10.1016/j.jempfin.2002.12.001.

Butler, K. C., & Malaikah, S. J. (1992). Efficiency and inefficiency in thinly traded stock markets: Kuwait and Saudi Arabia. *Journal of Banking and Finance, 16*(1). doi: 10.1016/0378-4266(92)90085-E.

Cohen, J. D. (2005). The vulcanization of the human brain: a neural perspective on interactions between cognition and emotion. *Journal of Economic Perspectives, 19*(4). doi: 10.1257/089533005775196750.

Cosma, G., & Acampora, G. (2016). *Neuro-fuzzy sentiment analysis for customer review rating prediction. Sentiment Analysis and Ontology Engineering: an environment of computational intelligence*. Springer International. doi: 10.1007/978-3-319-30319-2_15.

Fama, E. (1970). Efficient capital markets: a review of theory and empirical work. *Journal of Finance, 25*(2). doi: 10.1111/j.1540-6261.1970.tb00518.x.

Gallagher, L. A., & Taylor, M. P. (2002). Permanent and temporary components of stock prices: evidence from assessing macroeconomic shocks. *Southern Economic Journal, 69*(2). doi: 10.2307/1061676.

Gneiting, T., & Katzfuss, M. (2014). Probabilistic forecasting. *Annual Review of Statistics and Its Application, 1*. doi: 10.1146/annurev-statistics-062713-085831.

Gneiting, T., & Ranjan, R. (2013). Combining predictive distributions. *Electronic Journal of Statistics, 7*. doi: 10.1214/13-EJS823.

Haugen, R. A. (2001). *The inefficient stock market.* Pearson.

Heiden, S., Klein, C., & Zwergel, B. (2013). Beyond fundamentals: investor sentiment and exchange rate forecasting. *European Financial Management, 19*(3). doi: 10.1111/j.1468-036X.2010.00593.x.

Heston, S. L., & Sinha, N. R. (2015). News versus sentiment: predicting stock returns from news stories. *Robert H. Smith School Research Paper*.

Ho, K-Y., & Wang, W. W. (2016). Predicting stock price movements with news sentiment: an artificial neural network approach. *Artificial Neural Network Modelling. Studies in Computational Intelligence*, *628*. doi: 10.1007/978-3-319-28495-8_18.

Huang, C. L., & Goo, Y. J. (2008). Are happy investors likely to be overconfident? *Emerging Markets Finance and Trade, 44*(4). doi: 10.2753/REE1540-496X440403.

Jagongo, A., & Mutswenje, V. S. (2014). A survey of the factors influencing investment decisions: the case of individual investors at the NSE. *International Journal of Humanities and Social Science, 4*(4).

Kahneman, D., & Tversky, A. (1974). Judgment under uncertainty: heuristics and biases. *Science, 185*. doi: 10.1126/science.185.4157.1124.

Kahneman, D., & Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica, 47*(2). doi: 10.2307/1914185.

Kavussanos, M., & Dockery, E. (2001). A multivariate test for stock market efficiency: the case of ASE. *Applied Financial Economics, 11*(5). doi: 10.1080/09603100010013006.

Lucey, B. M., & Dowling, M. (2005). The role of feelings in investor decision making. *Journal of Economic Surveys*, *19*(2). doi: 10.1111/j.0950-0804.2005.00245.x.

Maghrebi, N., Holmes, M. J. & Oya, K. (2014). Financial instability and the short-term dynamics of volatility expectations. *Applied Financial Economics*, *24*(6). doi: 10.1080/09603107.2014.881966.

Maknickas, A., & Maknickienė, N. (2012). Influence of data orthogonally: on the accuracy and stability of financial market predictions. 4th International Joint Conference on Computational Intelligence. 5–7 October, 2012, Barcelona, Spain.

Maknickiene, N., & Maknickas, A. (2013a). Financial market prediction system with Evolino neural network and Delphi method. *Journal of Business Economics and management, 14*(2). doi: 10.3846/16111699.2012.729532.

Maknickiene, N., & Maknickas, A. (2013b). Investigation of prediction capabilities using RNN ensembles. 5th international joint conference on computational intelligence. 20–22 September 2013, Vilamoura, Portugal.

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance, 7*(1).

Oanda (2017). Data services for businesses and investors. Retrieved from https://www.oanda.com/.

Qian, B., & Rasheed, K. (2007). Stock market prediction with multiple classifiers. *Applied Intelligence, 26*. doi:10.1007/s10489-006-0001-7.

Plakandaras, V., Papadimitriou, T., Gogas, P., & Diamantaras, K. (2015). Market sentiment and exchange rate directional forecasting. *Algorithmic Finance*, *4*(1-2). doi: 10.3233/AF-150044.

Rechenthin, M., Street, W. N., & Srinivasan, P., (2013). Stock chatter: using stock sentiment to predict price direction. *Algorithmic Finance*, *2*(3-4). doi: 10.3233/AF-13025.

Rutkauskas, A. V., & Ramanauskas, T. (2009). Building an artificial stock market populated by reinforcement-learning agents. *Journal of Business Economics and Management, 4*(10). doi: 10.3846/1611-1699.2009.10.329-341.

Ryu, D., Kim, H. Y., & Yang, H. (2017). Investor sentiment, trading behavior, and stock returns. *Applied Economics Letters*, *24*. doi: 10.1080/13504851.2016.1231890.

Sahi, S. K. (2012). Neurofinance and investment behaviour. *Studies in Economics and Finance, 29*(4). doi: 10.1108/10867371211266900.

Seo, S. W., & Kim, J. S. (2015). The information content of option-implied information for volatility forecasting with investor sentiment. *Journal of Banking & Finance, 50*. doi: 10.1016/j.jbankfin.2014.09.010.

Sharda R., Delen D., & Turban E. (2013). *Business intelligence: a managerial perspective on analytics*. Prentice Hall Press Upper Saddle River.

Sheila, C. D. (2010). Cognition, market sentiment and financial instability. *Cambridge Journal of Economics, 35*. doi: 10.1093/cje/beq029.

Shu, H. C., & Chang, J. H. (2015). Investor sentiment and financial market volatility. *Journal of Behavioral Finance, 16*(3). doi: 10.1080/15427560.2015.1064930.

Simon, H. A. (1957). *A behavioral model of rational choice. Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*. New York, NY: Wiley.

Smales, L. A. (2017). The importance of fear: investor sentiment and stock market returns. *Applied Economics*, *49*(34). doi: 10.1080/00036846.2016.1259754.

Wang, B., Huang, Y., Yuan, Z., & Li, X. (2016). A multi-granularity fuzzy computing model for sentiment classification of Chinese reviews. *Journal of Intelligent & Fuzzy Systems, 30*(3). doi: 10.3233/IFS-151853.

Wang, T., Gong, Q., Ren, W., Wang, Y., Luo, X., & Li, Q. (2013). *A sentiment-based hybrid model for stock return forecasting.* 17th Pacific Asia Conference on Information Systems. 18–22 of June, 2013, Jeju Island, Korea.

Wurgler, J. (2007). Investor sentiment data (annual and monthly). Retrieved from http://people.stern.nyu.edu/jwurgler/.

Yang, H., Ryu, D., & Ryu, D. (2017). Investor sentiment, asset returns and firm characteristics: evidence from the Korean Stock Market. *Investment Analysts Journal*, *46*(2). doi: 10.1080/10293523.2016.1277850.

Annex

Table 1. Comparison of overall classification accuracy

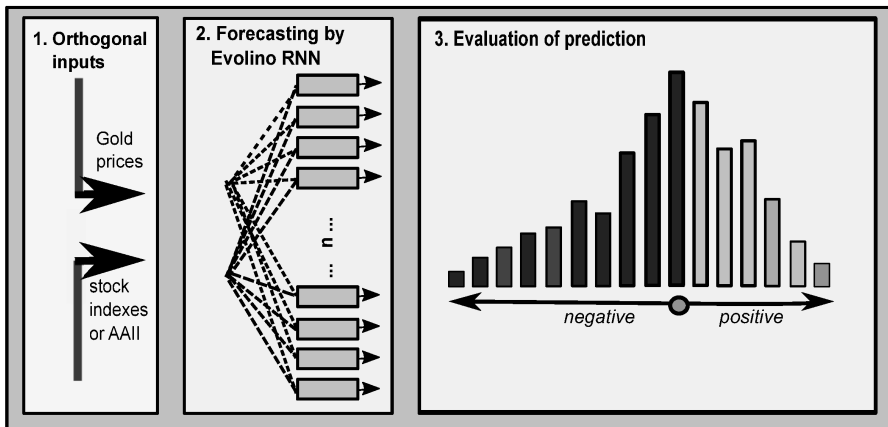| Sentiments		Finance market indices		Composition of AAII and stock indices all tests		Composition of AAII and stock indices, tests with high sentiments	
Number of predictions	OCA	Number of predictions	OCA	Number of predictions	OCA	Number of predictions	OCA
15	0.87	44	0.82	15	0.53	11	0.73

Figure 1. Scheme of the distribution-based model

Figure 2. Prediction of finance indexes a) increasing. b) decreasing. c) a risky state of proximally equal directions. d) when sentiments are significant
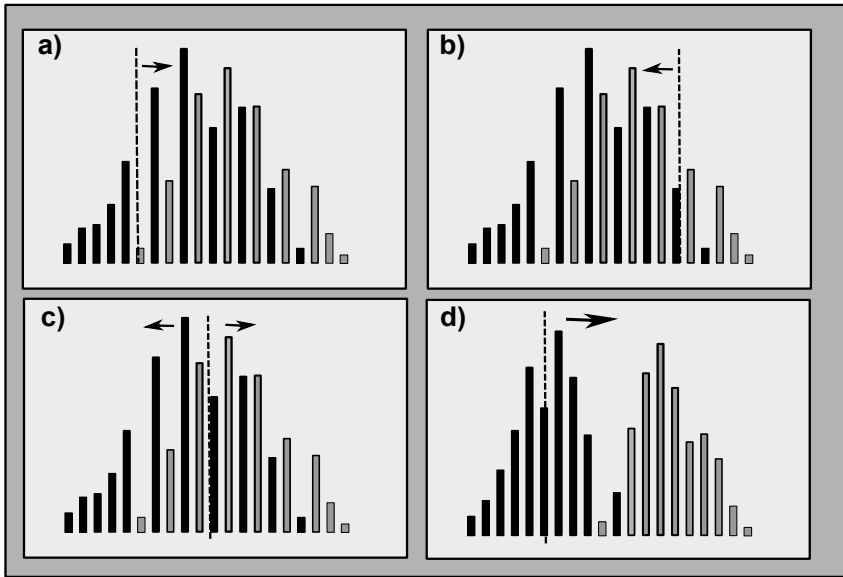
Figure 3. Dynamic of historical data for individual investors' sentiments (AAII) and the SPX500 index
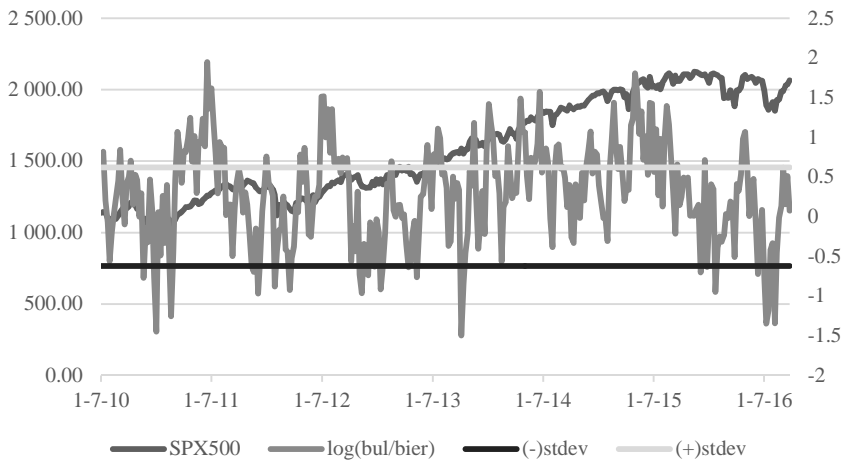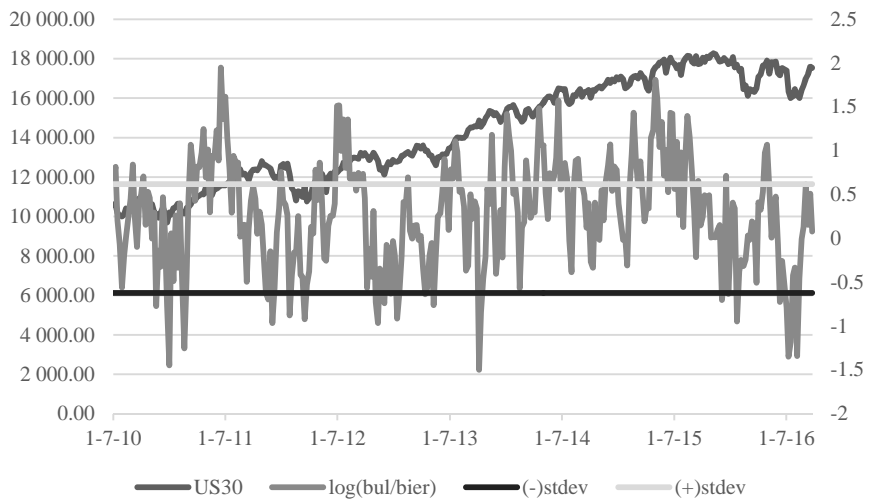

Source: own calculations based on Oanda (2017) and AAII (2017).

Figure 4. Dynamic of historical data for individual investors' sentiments (AAII) and the NAS100 index
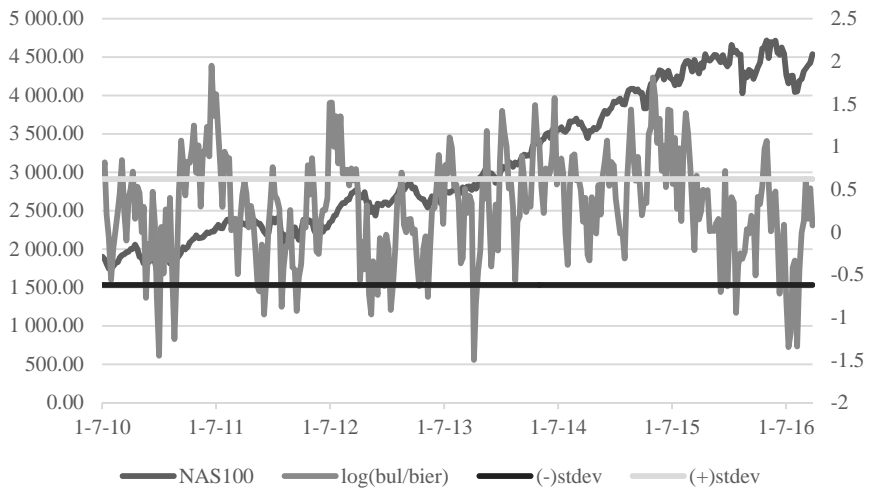
Source: own calculations based on Oanda (2017) and AAII (2017).

Figure 5. Dynamic of historical data for individual investors' sentiments (AAII) and the US30 index

Source: own calculations based on Oanda (2017) and AAII (2017).

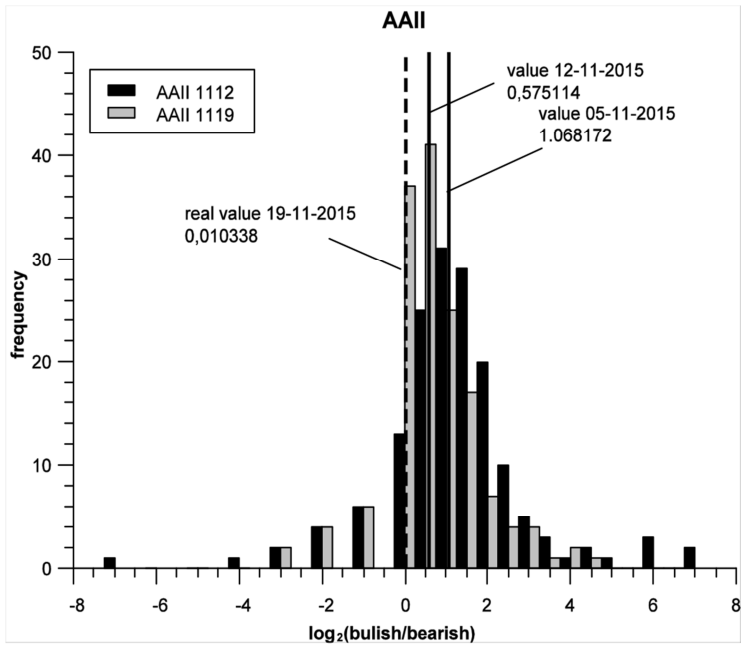**Figure 6.** Prediction of individual investors' sentiments (AAII)

Figure 7. Combination of the NAS100USD index and individual investors' sentiment (AAII) distributions
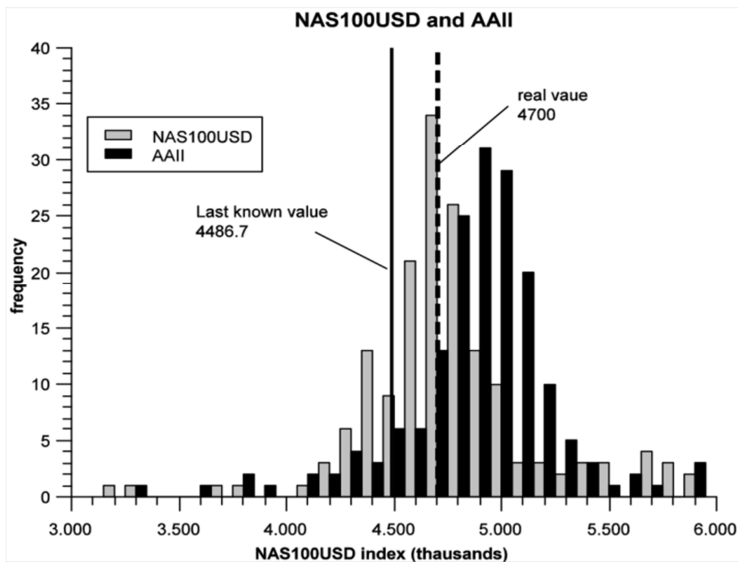
Figure 8. Combination of the US30USD index and individual investors' sentiment (AAII) distributions
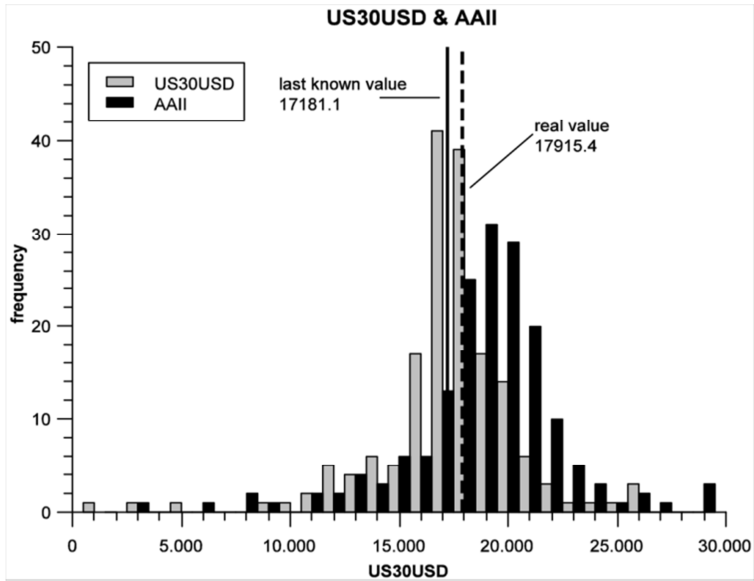

Figure 9. Combination of the SP500USD index and individual investors' sentiment (AAII) distributions