IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received September 28, 2020, accepted October 6, 2020, date of publication October 9, 2020, date of current version October 27, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3029860

An Efficient Word Embedding and Deep Learning Based Model to Forecast the Direction of Stock Exchange Market Using Twitter and Financial News Sites: A Case of Istanbul Stock Exchange (BIST 100)

ZEYNEP HİLAL KILIMCI[1] AND RAMAZAN DUVAR[2]
[1]Department of Information Systems Engineering, Faculty of Technology, Kocaeli University, 41000 Kocaeli, Turkey
[2]Department of Aircraft Electrical and Electronics, Faculty of Aeronautics and Astronautics, Kocaeli University, 41000 Kocaeli, Turkey

Corresponding author: Zeynep Hilal Kilimci (zeynep.kilimci@kocaeli.edu.tr)

ABSTRACT To forecast the movement directions of stocks, exchange rates, and stock markets are significant and an active research area for investors, analysts, and researchers. In this paper, word embedding and deep learning-based direction prediction of Istanbul Stock Exchange (BIST 100) is proposed by analyzing nine banking stocks with high volume in BIST 100. Though English news articles have been employed for forecasting of market direction previously, to the best of our knowledge, Turkish news articles and user comments from social media and different platforms have not been utilized with the combination of deep learning techniques and word embedding methods to predict the direction of Turkish stocks and market. For this objective, long short-term memory networks, recurrent neural networks, convolutional neural networks as deep learning algorithms and Word2Vec, GloVe, and FastText as word embedding models are evaluated. To demonstrate the effectiveness of proposed model, four different sources of Turkish news are collected. The news articles about stocks from Public Disclosure Platform (KAP), text-based technical analysis of each stock from Bigpara, user comments from both Twitter and Mynet Finans platforms are gathered. Experiment results demonstrate that the combination of deep learning techniques and word embedding methods have a great potential to predict the direction of BIST 100.

INDEX TERMS Deep learning techniques, financial sentiment analysis, stock market prediction, word embedding methods.

I. INTRODUCTION

The Istanbul Stock Exchange is expressed as BIST 100 index that is a frequently utilized abridgment for the Turkey's stock exchange. Istanbul Stock Exchange (ISE) includes the following price indices. These are the index of national all stocks, the index of national 30 stocks, the index of national 50 stocks, the index of national 100 stocks, the indices of sector and its sub sectors, the index of second national market, the index of novel economy market and the index of investment trusts. The index of national 100 stocks covers both the index of national 50 stocks and the index of national 30 shares and is employed as a fundamental index of the national market of it. Investors, analysts and researchers evaluate their investments, analysis, and researches with various parameters by performing fundamental or technical analysis of them in BIST 100. Fundamental analysis judges the value of an investment by investigating financial and economic elements while the movement direction of any investment is estimated using various parameters such as volume, price that are gathered from previous time slices of that investment. In this work, we focus on both fundamental and technical analysis are taken into account. For this purpose, different Turkish text resources are collected from different platforms to analyze the direction of movement of Istanbul Stock Exchange (BIST 100). The reason behind of focusing on fundamental

The associate editor coordinating the review of this manuscript and approving it for publication was Biju Issac.

and technical analysis is that some of investors can be influenced by comments made on various social media platforms and websites [1] and manage their own investments using this way. In particular, choosing a social media environment such as Twitter [2], [3] that keeps the pulse of the society and see the comments made in this environment as an investment tool for them can be considered normal for investors. Twitter, Bigpara, Public Disclosure Platform (KAP), and Mynet Finans are examples of different platforms that are analyzed in this work. Bigpara and Mynet Finans are websites that allow users to keep track of financial news and analysis. Moreover, Public Disclosure Platform (KAP) is an electronic system that informs the public in accordance with the capital market and change regulations. In this work, we propose an efficient word embedding and deep learning-based model to estimate the direction of BIST 100 with high volumed stocks in BIST 100. Through web crawler we developed, Turkish sources of financial text are collected from KAP, Mynet Finans and Bigpara websites. Selenium web browser [4] is used to collect individual and corporate comments from Twitter social media environment. After collecting text contents, the effect of dirty data is eliminated by using pre-processing techniques on the raw data sets. Then, data sets are labeled as positive and negative by using TextBlob [5]. After that, the performance of proposed model which is the consolidation of deep learning techniques, and word embedding methods is evaluated. To our knowledge, this is the first consideration to forecast BIST 100 direction employing the combination of deep learning techniques and word embedding algorithms on Turkish text-based fundamental and technical analysis. That is, the combinations of aforementioned deep learning and word embedding models are assessed for the first time in order to predict the direction of BIST 100 in this work. In order to assess the success of proposed method, nine bank stocks with a large transaction volume in BIST 100 are used. Experiment results reveal that our proposed model exhibits approximately 80% of accuracy result on Twitter, Mynet Finans, and KAP data sets. Moreover, to the best of our knowledge, this study is the primary endeavor to predict the stock market direction using the combination of deep learning algorithms and word embedding models on different Turkish text resources collected from different platforms.

Human activity recognition, speech recognition, and machine translation, image processing are research areas growing in popularity with the utilization deep learning techniques [6]–[9]. Newly, deep learning models are also proposed for text classification, and natural language processing (NLP) fields. Although deep learning is called as a sub category of machine learning, it usually presents more successful results than the conventional models because of its automatic feature extraction and deep neural networks properties. In addition to feature extraction, deep learning techniques are also used for classification purposes in many research areas. Long short-term memory networks (LSTMs) [6], [10], recurrent neural networks (RNNs) [11]–[13], convolutional neural networks (CNNs) [9] are well-known and applied architectures and frequently used in text categorization, and natural language processing problems. Moreover, the performance of machine learning techniques is mostly dependent on the representation of attributes for a problem. In natural language processing, and text mining fields, representation of attributes is rather considerable due to the large-scale number of words. Recently, learning task in the demonstration of attributes also gains popularity. This facilitates to learn demonstration of words from a text corpus which is named as word embedding. So, it can be summarized that the word embedding is a technique that learns the word representation from achieved corpus by learning models. Word2Vec [14], GloVe [15], FastText [16], BERT [17] as the methods of word embedding are frequently employed in natural language processing problems. Aforementioned deep learning methods and word embedding techniques are employed in this work for the aim of forecasting BIST 100 direction. Stock market prediction is a problem that acts trying to determine the future value of a company stock or other financial instrument traded on an exchange. For this reason, the successful prediction of a stock's future price could yield significant profit to its investors. On the other hand, sentiment classification task in stock exchange prediction is to forecast the direction of a company stock or other financial instrument as up or down without determining any value by analyzing opinions of persons. In this work, we focus on the direction prediction of Istanbul Stock Exchange by performing sentiment analysis task.

The rest of paper is presented as follows: In Section 2, some studies in the literature related to stock market forecasting are explained. Section 3 mentions on the proposed model and its details. Section 4 and 5 advert the experiment results and conclusion part, respectively.

II. LITERATURE REVIEW

In this section, a brief survey of literature works focused on the stock exchange forecast is introduced.

In the method proposed in [18], the stock prices of seven companies that make up the insurance sector index in the Istanbul Stock Exchange are estimated by artificial neural network models. In [19], a model established on the LSTM algorithm employing historical stock data is proposed to predict the direction of stocks on the Chinese stock exchange. They compare LSTM method with the random estimation model and conclude that the LSTM model exhibits more successful estimation for stock returns.

The Evolutionary Neural Network (ESA) is used in the method proposed in [7]. In this work, the prediction of status of stocks in BIST 100 are proposed by using ESA. The classification performance obtained by ESA is evaluated to be higher than those obtained with the chi-square feature selection and logistic regression classifier. The approach in [20] focuses on stock market prediction by evaluating deep learning methods for classification purpose. In [21], the accuracy rate of 94.21% prediction is obtained by using LSTM deep learning method in Turkish texts. In the method proposed in [22], data of 30 companies which are in Borsa Istanbul are utilized. In the study, authors propose to compare the results of ARIMA model as a time series analysis model with the results of LSTM algorithm as a deep learning technique. Experiment results show that ARIMA model outperforms LSTM. They also measure the performances of machine learning algorithms and conclude that linear regression gives better results than the other traditional machine learning techniques.

The approach presented in [23], the data of the 25 leading BIST 100 companies are analyzed and various forecasting algorithms are applied to data. In this work, it is shown that the random forest technique gives better experiment results. The presented method in [24], the movement direction of ISCTR, GARAN and THYAO stocks which are frequently traded over BIST 100 is estimated by utilizing deep learning methodology. For this purpose, convolutional networks are evaluated as a deep learning algorithm in the study and F-score measurement is employed in order to present the experiment results. Finally, accuracy results obtained with CNN are 0.574, 0.578, and 0.61 for the stocks ISCTR, THYAO, and GARAN, respectively.

In [25], a model for portfolio allocation is proposed. For this purpose, LSTM, multi-layer perceptron (MLP), and random forest classifier (RFC) are employed. Public financial sentiment and historical prices collected from the New York Stock Exchange (NYSE) are used to train machine learning models. In this work, it is shown that LSTM gives better experiment results. In [26], an approach is proposed to compute the asset-level market sentiment from social media data stream, and integrate it to the state-of-the-art asset allocation method using market views. Furthermore, it is proposed to use evolving clustering method (ECM) and LSTM approaches together to deal with noisy financial data.

The approach presented in [27], the authors proposed to combine the technical and fundamental analysts approaches to market trend forecasting through the use of conventional machine learning techniques applied to time series prediction and sentiment analysis. In [28], experiments have been conducted on more than five years of real Hong Kong stock market data using four different sentiment dictionaries. In this work, a two-layer LSTM network is proposed to learn the sequential information within the series of the indicators and news sentiments, together with a fully connected network that makes final stock price trend predictions. Multiple kernel learning (MKL) and support vector machine (SVM), are employed as benchmarks to compare the performances of proposed approach and it is shown that the proposed approach outperforms the baseline methods.

The method presented in [29], a convolutional neural network-based classifier is proposed to extract users' tendencies from their comments, and introduce the distributed lag model and the GARCH model to investigate the impact of users' tendencies on market volatility and market returns. In [30], it is proposed to use deep convolutional networks to estimate the stocks' direction traded on the US NASDAQ Stock Exchange. The success of proposed model is appraised with evaluation metrics namely, F-criterion, and accuracy. Eventually, authors conclude the study that proposed technique presents the best success with Deep Convolutional Neural Networks (DVM) method.

Our study differs in literature studies mentioned above in that it employs user reviews, announcements, and Turkish text-based financial technical analyses about stocks from both social media platform and financial websites. Moreover, the novelty side of this study is the usage of combination of word embedding techniques and deep learning models for forecasting the direction of BIST 100. In addition, we propose to provide a valuable text-based financial resource to guide investors' forecasting of stock quotes using new generation models.

III. METHODOLOGY

A summary of the methods, materials, and proposed framework are introduced in this section.

A. WORD EMBEDDING MODELS

Word embedding is known as a feature learning and language modeling technique. Word vectors are created using word embedding models. For example, currencies such as "dollar", "euro" and "pound" are placed close to their semantics, but the word "monkey" will be located far from these words in the word vector. In this work, Word2Vec, FastText and GloVe word embedding models are utilized.

1) Word2Vec

Word2Vec is accepted as a pioneer word embedding method that starts a new trend in natural language processing. Word2Vec tries to express words in a vector space and it is a prediction-based and unsupervised model [8]. Thanks to neural networks, the model can learn easily representation of words as dense vectors that encode patterns and many linguistic regularities among words. Thus, it makes possible to display trained words as vectors, encoding multiple language models between words. There are 2 types of sub-methods, Skip-Gram and CBOW (Continuous Bag of Words). Although both methods are generally similar, they have different advantages compared to each other.

The purpose of the skip-gram model is to predict the words surrounding a given word. Given a center word and sequence of training words $w_1, w_2, w_3, \ldots, w_t$ skip-gram model maximizes the average log probability of n surrounding words of the center word $\mathrm{w_t}$:

$$J_\theta = \frac{1}{T}\sum_{t=1}^{T}\sum_{-n\leq j\leq n} log\, p\left(w_{t+j}\middle| w_t\right) \tag{1}$$

In this equation n denotes the size of training context. CBOW model predicts a target word w_t from the surrounding words by maximizing the log probabilities.

$$J_\theta = \frac{1}{T}\sum\nolimits_{t=1}^{T} log\, p\left(w_t\middle| w_{t-n}, \ldots, w_{t-1}, w_{t+1}, \ldots, w_{t+n}\right) \tag{2}$$

Given n words before and the after the target word w_t, the CBOW learns the target word w_t from a training corpus and later it can be used to predict a target word w_t from surrounding words.

2) GloVe

Global Vectors (GloVe) is another popular word embedding algorithm which is introduced in [9].

Word2Vec models use surroundings of words for training and do not take advantage of the count-based statistics which includes word co-occurrences. For this purpose, GloVe method consolidates the local content window and count-based matrix factorization techniques in order to achieve more effective representation. Matrix factorization provides obtaining word to word statistical information from a corpus.

Glove method first contracts a word co-occurrence matrix X. Each element of X_{ij} shows the number of times word i appears in the context word j. The Glove model utilizes (3) to calculate cost.

$$J_\theta = \sum_{i,j=1}^{V} f\left(X_{ij}\right)\left(w_i^T w_j + b_i + b_j - logX_{ij}\right)^2 \tag{3}$$

In this equation, w_i the vector for the main word, w_j is the vector for context word and, b_i, b_j are scalar biases for the main and context words.

3) FastText

FastText is an artificial neural network library developed for text classification. Converts text or words into continuous vectors that can be used in any language, such as a speech-related task. In this context, detection of spam can be one of the most common examples. It is faster and more efficient than other text classification structures. Instead of using individual words as inputs, it divides the words into several letter-based "n-grams". N is the n repetition degree in gram expression. The word is divided into characters with the expression n, which allows us to understand how much the length of a word is [16], [31]. FastText uses the skip-gram model with negative sampling proposed for Word2Vec with a modified skip-gram loss function. Let $G_w \supset \{1, \ldots, G$ be the set of n-grams appearing in a word w, the score of the word is calculated by the sum of the vector representations of its n-grams:

$$s(w, c) = \sum\nolimits_{g \in G_w} z_g^T v_c \tag{4}$$

4) BERT

Bidirectional Encoder Representations from Transformers (BERT) is a machine learning framework which is designed for natural language processing. BERT is a deep learning model that every output element is connected to every input element and it is designed to pre-train deep bidirectional representations from unlabeled text by jointly conditioning on both left and right context in all layers. In this framework, weightings between elements dynamically calculated based upon their connection. In this study, BerTurk version is employed which is a community-driven cased BERT model for Turkish.

B. DEEP LEARNING MODELS

In this part, Long Short-Term Memory Networks (LSTMs), Recurrent Neural Networks (RNNs), and Convolutional Neural Networks (CNNs) are introduced.

1) CONVOLUTIONAL NEURAL NETWORKS

CNNs are very successful in the area of image processing, and in addition during the recent years, it is observed that they are also successful in NLP problems [32]–[34]. CNN is known as a feed forward neural network which includes pooling, convolution, and full connected layers. There can be many convolution layers that performs a filter of convolution to data in order to acquire features which are fed into pooling layers and followed by dense layers. The fundamental task of filters is to learn the context of problem throughout training procedure. In this way, dependencies located in original data is represented with utilization of feature maps which is named convolution process. Then, the pooling layer is used to decrease the parameters and the number of calculations in the network with the purpose of decreasing training time and reducing dimension and over-fitting. After that, the final decision is assigned by fully connected layers.

2) RECURRENT NEURAL NETWORKS

In RNNs, the output from the preceding step feeds the current step as input to remember words. RNN use a hidden layer to remember information which are calculated in the past. RNN, as distinct from other neural networks, reduces the semantic difficulty of inputs to be set. RNN applies the same operations on all inputs or covered layers to produce the result. Using the same data for each input decrease the semantic difficulty of the data [11]–[13]. When long-term dependencies are seen in sequence data, RNN based models cannot learn previous data, properly. The reason of this problems are gradient descent operations performed in back-propagation process. As a result of continuous matrix multiplications, small weight values decrease exponentially and disappear. Moreover, when the weight values are large, these values reach "NaN" values as a result of the continuous matrix multiplication. To handle these kinds of issues, techniques such as suitable activation functions gradient or clipping can be utilized.

3) LONG SHORT-TERM MEMORY NETWORKS

LSTMs are advanced to handle gradient based problems of RNNs. They are sub-branch of RNNs which can maintain information in memory for long periods of time. Thus, long dependencies among data is stored and the contextual semantics are kept with the usage of LSTMs. The starting point is to ensure a solution to the exponential error growth problem using back propagation algorithm while deep neural networks are being trained. Errors are stored and used by LSTMs in

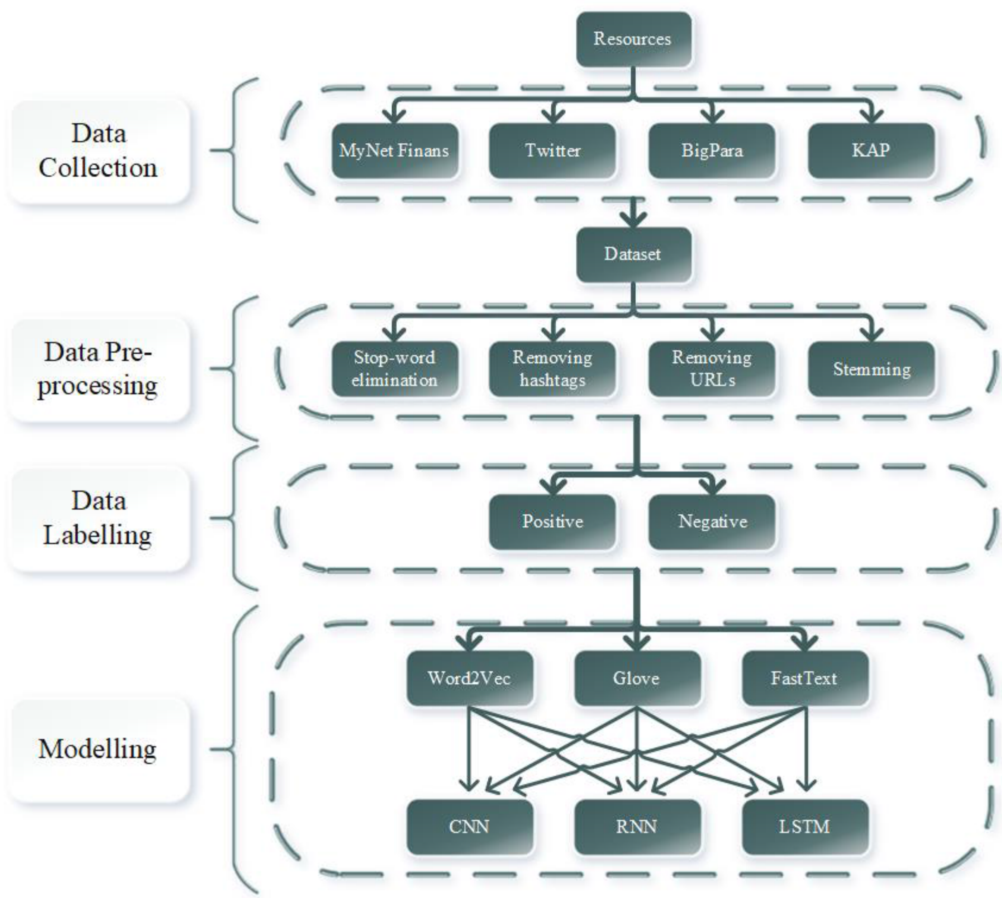

FIGURE 1. A general flow chart of the proposed model.

back propagation process. Decisions can be made by LSTM, such as what to keep and when to authorize reads [10], [35].

C. PROPOSED FRAMEWORK

In this work, the movements of banking stocks with the highest transaction volume in BIST 100 are used. Since these stocks have high trading volume, it is considered that there is a connection between the movements of these stocks and the direction of the stock market. A general flow chart is given in Fig. 1.

1) DATA PREPARATION

In the experiments, we employ four different data sets namely, Twitter, Bigpara, Public Disclosure Platform (KAP), and Mynet Finans. Twitter is a social networking service on which users post and interact with messages named as tweets that are restricted to 280 characters. In this way, the Twitter dataset, which we also used in our study, is a platform that allows users to analyze their opinions about any topic, product or service. Mynet is a local internet platform with more than 30 content services that try to meet the demands of its users with content ranging from news to video, entertainment to education. Mynet Finans service used in this study provides its users with stock market data instantly and live. It provides a platform with all analysis, comments and data related to stock market, investment, economy and money from the world and Turkey. Finally, it is an environment that offers instant news from domestic and global markets that guide investors ' investments and provides international market data to its users instantly. BigPara is a subsidiary of Hurriyet Inc.. In addition to providing information about financial markets, Bigpara allows its users to make credit comparisons. In this way, instead of comparing individual interest rates, total repayments, installment amounts, and expenses from many banks' websites, it offers the opportunity to do all of this from the BigPara platform. It offers a compact environment where its users can follow instant stock prices, dollar, euro and other exchange rates, gold and oil prices, current news. The BigPara data set, which is also used in our study, evaluates the technical analysis of experts on this platform and provides a text-based analysis report. KAP is an electronic system in which transactions and notifications that need to be disclosed to the public in accordance with the capital market and Istanbul Stock Exchange legislation are transmitted as e-signatures. In general, large subsidiaries, stock holders and partners traded in Istanbul Stock Exchange, brokerage firms, investment companies authorized by the Capital Market Board and various funds can send direct notifications to KAP. On this platform, notifications can be viewed online. The language of the notifications sent is Turkish, but the company's balance sheets are also published in English.

At a first stage of this work, a browser program is developed to collect data from websites. This program gathers data through a scan-to-scan approach using keywords and direct link addresses. User comments from the Mynet Finans website, text-based technical analysis comments from the Bigpara website, annotations and news from the Public Disclosure Platform (KAP) are gathered for related stocks. The following keywords of the stocks are employed during the data preparation process: "AKBNK", "ALBRK", "GARAN", "HALKB", "ISCTR", "SKBNK", "TSKB", "VAKBN", "YKBNK". Similarly, user opinions from Twitter environment are gathered utilizing keywords and hashtags through the Selenium browser at the next step. The data sets obtained from Bigpara, Mynet Finans, Public Disclosure Platform (KAP), and Twitter are generated from the data which is located between September 1, 2018 and September 1, 2019. Since there are daily comments that are interpreted by analysts by performing text-based technical analysis for different stocks in BIST 100 on this site, historical data of stocks cannot be collected. Because of this reason, text-based technical analysis from Bigpara platform are gathered daily between September 1, 2018 and September 1, 2019.

Since the initial form of the data sets is very dirty, the data sets are cleaned through pre-processing methods at this stage. The pre-processing step that includes stop-word elimination, removing hashtags and removing URLs is carried out using pandas, NumPy and regex libraries. Stemming as another preprocessing model is performed by TurkishStemmer library. Since the supervised machine learning techniques are employed in this study, the collected Turkish texts are labeled. For the purpose of labeling the Turkish data sets cleaned with pre-processing techniques, the data set are trained with the pre-trained "Hepsiburada" data set which is downloaded from Kaggle [36]. At this step, naive Bayes classifier in TextBlob is utilized to demonstrate the sentiments by tagging each document as negative or positive. While it is possible to perform sentiment analysis on English texts using the labeled data set provided by TextBlob for pre-training, since there is no labelled data set for the Turkish language, it is proposed to train the Turkish data sets through labeled Hepsiburada data set which is shared on Kaggle. Hepsiburada data set has two columns, which include user reviews and scores, respectively. In Hepsiburada data set, documents are rated as 1, 2, 3, 4, and 5. In this study, 4 and 5 scores mean that the sentiment of the document is positive and 1 and 2 scores are assigned as negative sentiment. The statistics achieved after pre-processing and labeling steps are given in Table 1. The content of each data set is presented in Table 2.

TABLE 1. The statistics of the data sets.

Dataset	Positive	Negative	Total	Avg. #of terms per doc	Avg. term length
Twitter	9,075	7,812	16,887	18,22	5,37
Mynet Finans	5,775	5,001	10,776	24,70	6,13
KAP	4,110	3,457	7,567	175,52	7,42
Bigpara	4,098	2,850	6,948	205,41	6,68

TABLE 2. Contents of the data sets.

Dataset	Content
Twitter	akbnk 5.44 desteğinden başlayıp 5.90 direncine kadar gidecek.
Mynet Finans	buradan satılırmı? yorumlarda 7-7,10 aralığı satılabilir deniliyordu 7 üstünede çıkarmı
KAP	AKBANK T. A.Ş.'nin sermayesinin 4.000.000.000 TL'den 5.200.000.000 TL'ye artırılması sebebiyle nakit karşılığı ihraç olunan 1.200.000.000 TL pay senedinin...
Bigpara	AKBNK hissesi günü 5,87 TL'den ve 1,73 deger kazancı ile tamamladı. Hisse senedi kapanış ile 10 günlük hareketli ortalamaların üzerinde bulunuyor. Hisse senedinin kısa ...

2) WORD EMBEDDING PROCESS

The data set that is ready for modeling is sent as an input to word embedding models at the first stage, and the word vectors derived from word embedding models are forwarded as input to deep learning algorithms. Before proceeding to classification stage, 100-dimensional word vectors are constructed for each word in a document. The similarity score of each word in the document with the selected term is calculated. Then, the output of each word embedding model is fed into the layers of deep learning models as an input. Thus, we extract/select the features represented with dense vectors from the huge corpus to send deep learning models as input. After that, classification is carried out using word embedding and deep learning models on these data sets, which are ready after pre-processing and tagging processes.

In classification step, the proposed hybrid approaches are presented by combining aforementioned word embedding models and deep learning algorithms. Word embedding models namely, FastText, GloVe, Word2Vec, deep learning techniques such as CNN, LSTM, RNN as well as combinations of these methods are approaches that are evaluated in this study. For the deep learning models and word embedding algorithms, Keras library is utilized. We carry out too many experiments in order to measure the impact of parameters such as batch sizes, window sizes, training set sizes, epoch sizes. After arranging parameter settings by changing epoch size, batch size, window size, and training set size, the best classification accuracies are presented in this study. Firstly, we focus on the effect of training set percentage by changing training set sizes 80, 50, and 30. The best classification results are obtained when training set percentage is set to 80. After varying epoch size between 1 and 5, almost the same or a decrease down to 2 percent is observed in trials where epoch size is greater than 2. Because of this, epoch size is set to 2. Then, the best classification accuracies are attained with batch size 64 where it ranges from 32 to 256. For all experiments we also investigate the impact of windows size values that is diversified as 1,3,5,7. The best performance is exhibited when window size is set to 1 compared to the others. The experiment results exhibited in Section IV are obtained when size is 100, window is 1, min_count is 3, workers are 3, and sg parameter is set to 1 for Word2Vec model that means Skip-gram model is evaluated.

3) DEEP LEARNING PROCESS

After deriving word vectors, the proposed approach is ready for modelling. Then, nine different hybrid deep learning models are constructed, Word2Vec+CNN, GloVe+CNN, FastText+CNN, Word2Vec+RNN, GloVe+RNN, FastText+RNN, Word2Vec+LSTM, GloVe+LSTM, FastText+LSTM, where each word embedding model is used as input to deep learning methods. As a result of all these stages, the model with the best classification success is selected to forecast the direction of Istanbul Stock Exchange.

Furthermore, the deep learning networks used in this work consist of three layers. In the CNN technique, the number of filters in the first layer is adjust to 128 and the core size is 5. Rectified linear unit (ReLU) is selected as the function of activation. In the second layer, the maximum pool size is determined as 2 and GlobalMaxPooling is used. In the third layer, the dense full connected layer is set as the sigmoid activation function. In the RNN deep learning model, LSTM 64 is set by selecting bidirectional in the first layer. The drop layer is used to overcome over-fitting in the second layer and adjusted to 0.2. In the third layer, the activation function for the dense fully connected layer is determined as sigmoid. For the LSTM model, a layer named FC1 with 256 outputs is created. Then, the activation function in the second layer is designated as ReLU. In the third layer, the drop layer is used to prevent over-fitting. In order to avoid over-fitting especially on smaller size of data sets like Bigpara, we try many methods like dropout, regularization techniques, early stopping, simplying model, weight constraint. In our study, we observe that the utilization of early stopping with dropout and a weight constraint are more effective method in order to prevent over-fitting. 0.2 is obtained as an optimum value for dropout among 0.3, 0.4, 0.5, 0.6. Epoch size is set to 2 as an early stopping criteria and weight constraint is adjusted to 4.

IV. EXPERIMENTAL RESULTS

In this work, far-reaching experiments are accomplished with the combination word embedding and deep learning techniques for the purpose of forecasting the direction of BIST 100 by using high volume stock market shares in BIST 100. Data sets include user reviews, text-based technical financial analyzes, and financial news from the sites determined about the large volume stock market shares in BIST 100. Accuracy is employed as an evaluation metric in the experiments to demonstrate the classification success of each model and the contribution of our study. All models are run on Google colab environment provided free GPU usage by Google.

Experiments are accomplished using 80% training and 20% test of data with repeated holdout method. It is applied 10 times on each data set. This method is similar to the previous works [37]–[39] where they use the same split with training and test data. The following abbreviations are employed for the combination of word embedding techniques and deep learning methods: W+C: Combination of Word2Vec and CNN, W+R: Combination of Word2Vec and RNN, W+L: Combination of Word2Vec and LSTM, G+C: Combination of GloVe and CNN, G+R: Combination of GloVe and RNN, G+L: Combination of GloVe and LSTM, F+C: Consolidation of FastText and CNN, F +R: Combination of FastText and RNN, F+L: Consolidation of FastText and LSTM. The combination is interpreted as vectors originating from Word2Vec, GloVe, and FastText word embedding models are given as an input into CNN, RNN and LSTM algorithms and an accuracy value is generated. The best accuracy percentages are represented in bold letters. The results of these accuracy values are given in Table 3, Table 4, Table 5 and Table 6, respectively.

In Table 3, the performance of combination of deep learning techniques and word embedding methods in terms of each bank stock on Twitter data set is given. The combination of Word2Vec and LSTM outperforms other hybrid models with 80.61% accuracy value when average classification success is considered. G+L combination and all deep learning combinations of FastText model exhibit close performance to W+L with approximately 80% success rate. It is clearly observed that the selection of LSTM model as a deep learning technique generally improves the performance of system. Although the combination of W+L seems to work as the best, the combination of the FastText model with all deep learning algorithms is generally higher performance than the success rate of other combinations. The combinations of FastText model with deep learning models are 80.37%, 80.13% and 80.23%, respectively. Therefore, since the performance of W+L with 80.61% is very close to all combination results of FastText, it can be said that all deep learning algorithms modeled with FastText in the Twitter data set generally give much better results than the other combinations. When the performance of proposed model is considered as stock-based for AKBNK, the utilization of GloVe as a word embedding model outperforms other combinations for each deep learning algorithms. In AKBNK stock, LSTM also presents significant contribution to each word embedding model for the purpose of constructing combinations.

In Table 4, accuracy results of the combination of deep learning algorithms and word embedding models on Mynet Finans data set for each bank stock is presented. The combination of FastText and LSTM with 86.39% accuracy result notably performs well compared to the other models. 85.56% of accuracy for F+R and 84.81% of accuracy for W+L are observed, respectively. The mean classification result is obtained about 78% of accuracy when the consolidation of

TABLE 3. Accuracy results of the combination of deep learning algorithms and word embedding models on Twitter data set for each bank stock.

Twitter	Models									
Stocks	W+C	W+R	W+L	G+C	G+R	G+L	F+C	F+R	F+L	BerTurk
AKBNK	72.46	83.94	87.25	86.53	88.68	88.44	84.28	84.87	88.36	86.47
ALBRK	88.54	87.47	89.46	86.98	87.99	81.68	89.39	85.36	82.33	89.56
GARAN	87.23	85.27	78.98	85.75	81.51	80.72	82.48	80.59	78.83	80.65
HALKB	82.46	73.94	77.25	76.53	78.68	78.44	74.28	74.87	78.36	79.25
ISCTR	76.54	77.47	79.46	76.98	77.99	81.68	79.39	85.36	72.33	80.38
SKBNK	75.54	78.47	79.46	75.98	78.59	82.68	78.39	75.36	82.33	81.23
TSKB	72.46	73.92	75.20	56.53	58.68	68.44	74.21	74.85	78.34	80.74
VAKBN	75.54	76.47	79.46	77.98	78.59	80.68	78.39	79.36	82.33	81.55
YKBNK	81.23	80.87	78.98	85.75	81.51	80.72	82.48	80.59	78.83	80.39
Avg.	79.11	79.76	**80.61**	78.78	79.14	80.39	80.37	80.13	80.23	82.25

TABLE 4. Accuracy results of the combination of deep learning algorithms and word embedding models on Mynet Finans data set for each bank stock.

Mynet Finans	Models									
Stocks	W+C	W+R	W+L	G+C	G+R	G+L	F+C	F+R	F+L	BerTurk
AKBNK	87.98	88.59	70.68	88.39	89.36	72.33	84.21	84.85	89.94	89.05
ALBRK	84.67	84.77	88.85	82.66	83.82	85.20	86.53	86.68	89.56	90.57
GARAN	84.25	82.75	84.81	78.31	78.79	79.74	80.83	85.56	88.15	88.24
HALKB	77.98	78.59	80.68	78.39	79.36	82.33	74.21	74.85	85.57	81.15
ISCTR	74.67	74.77	78.85	72.66	73.82	75.20	76.53	76.68	80.65	80.95
SKBNK	74.61	74.75	78.84	72.66	73.82	75.20	76.53	78.68	81.39	81.57
TSKB	77.98	78.59	80.68	78.39	79.36	82.33	74.21	74.85	83.03	81.91
VAKBN	87.98	88.59	70.68	88.39	89.36	72.33	84.21	84.85	89.23	89.23
YKBNK	84.67	84.77	88.85	82.66	83.82	85.20	86.53	86.68	90.05	91.02
Avg.	84.25	82.75	84.81	78.31	78.79	79.74	80.83	85.56	**86.39**	85.97

GloVe models are considered with each deep learning model. Thus, the utilization of GloVe model for demonstration of documents drops the classification success to the poorest accuracy value.

In Table 5, the classification accuracies of the mixture of deep learning techniques and word embedding methods on Bigpara dataset for each stock is demonstrated. The consolidation of Word2Vec and RNN with 77.82% accuracy value significantly outperforms the other models. It is followed by G+R with 76.15% of accuracy and W+C with 75.85% of accuracy results, respectively. This means the power of W+R boosts the system performance by providing nearly 9% improvement compared to the poorest classification success of G+C. Furthermore, the performance enhancement of W+R in GARAN stock reaches roughly 8% compared to the combined model W+C. Moreover, it is clearly seen that RNN boosts the system performance when it is combined with any word embedding model. Thus, the usage of RNN as a deep learning model is more suitable especially combined with Word2Vec in terms of overall performance. The poorest classification success is exhibited when RNN is consolidated by FastText with 71.73% of accuracy. As a result of Table 5, W+R is the best model with 77.82% of accuracy result when mean classification results is considered.

The classification performance of combination of deep learning techniques and word embedding methods in terms of each bank stock is introduced on KAP data set in Table 6. Although GloVe and FastText combinations exhibit close classification performances, the blend of FastText and RNN with 79.74% accuracy slightly outperforms GloVe and RNN. It is clearly observed that F+R provides almost 2% improvement compared to the poorest combinations of Word2Vec. Considering that RNN is the deep learning model that offers the best success, it is observed that the poorest combination is performed by Word2Vec model and the most successful consolidation is achieved by FastText model in terms of average accuracy results. As a result of Table 6, the combination of FastText and RNN is the best consolidated model. Due to LSTM is the customized version of RNN, these exhibit similar classification performances in the experiments. It is observed in Twitter, Mynet Finans, and KAP datasets with comprehensive document size that FastText model contributes to the classification performance, significantly. In summary, when the experimental results are evaluated, it is observed that the combination of FastText and LSTM generally offers remarkable results when the data set is of sufficient size.

Moreover, the performance of proposed hybrid model for each data set is presented in Tables 7, 8, 9, and 10 to compare with word embedding models and deep learning methods. In Table 7, the combination of Word2Vec and LSTM exhibits the best classification performance with 80.61% of accuracy on Twitter data set. It is followed by LSTM model and Word2Vec method with 79.80% of accuracy value and 78.92% of accuracy result, respectively. In Table 8, FastText+LSTM consolidation slightly outperforms baseline models namely, FastText and LSTM on Mynet Finans data set. FastText and LSTM exhibit almost similar classification performance with 85.69% and 85.99% of accuracies, respectively. In Table 9, the hybrid model Word2Vec+RNN

TABLE 5. Accuracy results of the combination of deep learning algorithms and word embedding models on Bigpara data set for each bank stock.

BigPara	Models									
Stocks	W+C	W+R	W+L	G+C	G+R	G+L	F+C	F+R	F+L	BerTurk
AKBNK	80.00	80.02	74.13	78.00	77.02	77.13	70.00	70.02	62.13	79.15
ALBRK	82.02	80.07	75.12	72.02	75.07	71.12	72.02	71.07	70.12	78.90
GARAN	72.87	81.18	77.43	69.87	73.18	80.03	73.87	72.18	72.43	77.27
HALKB	74.00	80.11	60.01	66.00	85.11	74.01	72.00	77.11	71.01	76.06
ISCTR	76.43	77.00	75.31	65.43	79.00	75.31	69.43	66.00	77.31	75.44
SKBNK	75.59	73.87	74.57	67.59	75.87	73.12	71.14	75.54	63.75	71.96
TSKB	72.21	75.57	69.56	63.21	73.57	71.56	71.21	78.57	71.56	73.28
VAKBN	74.59	76.96	68.20	72.59	73.96	72.50	68.19	67.56	72.85	74.53
YKBNK	74.90	75.56	67.97	61.90	72.56	70.17	67.90	67.56	73.15	72.40
Avg.	75.85	**77.82**	71.37	68.51	76.15	73.88	70.64	71.73	70.48	75.44

TABLE 6. Accuracy results of the combination of deep learning algorithms and word embedding models on KAP data set for each bank stock.

KAP	Models									
Stocks	W+C	W+R	W+L	G+C	G+R	G+L	F+C	F+R	F+L	BerTurk
AKBNK	84.87	82.34	85.89	85.27	85.49	81.80	86.67	88.48	84.69	87.36
ALBRK	84.87	82.34	85.89	85.27	85.49	81.80	86.67	83.48	84.69	81.55
GARAN	81.23	80.87	78.98	87.67	82.25	81.86	78.36	76.79	75.69	75.52
HALKB	74.87	72.34	75.89	75.27	75.49	71.80	76.67	73.48	74.69	71.42
ISCTR	74.21	74.85	78.34	72.46	73.92	75.20	76.53	78.68	78.44	76.10
SKBNK	82.21	86.57	71.56	72.21	83.57	81.56	72.21	88.57	81.56	84.66
TSKB	78.56	75.42	75.33	74.21	74.85	78.34	78.99	77.89	79.73	75.48
VAKBN	64.87	62.34	65.89	75.20	71.49	71.90	76.64	73.46	74.62	70.90
YKBNK	81.23	80.87	78.98	87.67	82.25	81.86	78.36	76.79	75.69	74.67
Avg.	78.55	77.55	77.42	79.47	79.42	78.46	79.01	**79.74**	78.87	77.52

TABLE 7. Comparison with the best combined models, deep learning methods and word embedding techniques on twitter data set for each bank stock.

Stocks	Word2Vec	LSTM	Word2Vec+LSTM
AKBNK	84.17	85.39	87.25
ALBRK	87.23	88.05	89.46
GARAN	76.65	78.23	78.98
HALKB	75.96	76.85	77.25
ISCTR	77.16	77.54	79.46
SKBNK	78.57	79.45	79.46
TSKB	74.72	75.93	75.20
VAKBN	78.80	78.64	79.46
YKBNK	77.05	78.12	78.98
Avg.	78.92	79.80	**80.61**

enhance the classification performance of the system nearly 2% improvement in classification accuracy on BigPara data set compared to the Word2Vec and RNN models, separately. In Table 10, it is clearly observed that the combination of FastText and RNN model demonstrates remarkable classification results with 79.74% of accuracy on KAP data set. It is followed by FastText and RNN models, respectively. As a result of Tables 7, 8, 9, and 10, instead of using separately baseline methods such as word embedding techniques and deep learning models, combination of both methodologies exhibits superior classification performance.

As a consequence of all tables, we observe the superior classification success of recurrent neural networks, and long -short term memory networks that is specialized form of RNN compared to the CNN model for all data sets. CNN could not exhibit an extraordinary classification performance because feature extraction step that is the powerful side of CNN algorithm is already implemented by word embedding models. Moreover, the experiment results are consistent through the structure of RNN and LSTM models. It is not surprise that the superior performance of RNN and LSTM because of providing long-short word dependencies when considering the CNN architecture. Thus, the combination of RNN and LSTM models with word embedding models exhibit remarkable experiment results for all data sets. When data sets are evaluated, it is observed that either Word2Vec or FastText word embedding models present more successful results. Mynet Finans and Twitter data sets contain user comments which have much shorter text content compared to the Bigpara and KAP data sets. So, it is clearly seen that LSTM achieves short term dependencies between words in Mynet Finans and Twitter data sets even though RNN exhibits very close classification performance. Furthermore, compared to Twitter in terms of length of user comments, Mynet Finans dataset has longer text content. This is also similar to Bigpara and KAP data sets, where the KAP data set has longer text content. It is observed that FastText model contributes about 2% more to classification success in data sets with higher text lengths. Here, we have seen that the usage of FastText model at the feature extraction phase of the longer data sets based on the text lengths in datasets with similar contents contributes to the classification performance.

We think this is due to the fact that the words which can be learned with word-level trained models such as Word2Vec

TABLE 8. Comparison with the best combined models, deep learning methods and word embedding techniques on Mynet Finans data set for each bank stock.

Stocks	FastText	LSTM	FastText+LSTM
AKBNK	89.55	89.23	89.94
ALBRK	89.27	89.00	89.56
GARAN	88.75	88.72	88.15
HALKB	85.43	85.15	85.57
ISCTR	80.27	80.72	80.65
SKBNK	80.78	80.66	81.39
TSKB	81.24	82.80	83.03
VAKBN	87.90	88.24	89.23
YKBNK	88.00	89.37	90.05
Avg.	85.69	85.99	**86.39**

TABLE 9. Comparison with the best combined models, deep learning methods and word embedding techniques on BigPara data set for each bank stock.

Stocks	Word2Vec	RNN	Word2Vec+RNN
AKBNK	77,25	78,33	80,02
ALBRK	77,10	78,40	80,07
GARAN	78,42	78,77	81,18
HALKB	78,77	78,90	80,11
ISCTR	75,23	75,27	77,00
SKBNK	71,36	72,34	73,87
TSKB	74,30	74,06	75,57
VAKBN	75,50	75,25	76,96
YKBNK	74,80	74,00	75,56
Avg.	75,86	76,15	**77,82**

TABLE 10. Comparison with the best combined models, deep learning methods and word embedding techniques on Twitter data set for each bank stock.

Stocks	FASTTEXT	RNN	FastText+RNN
AKBNK	87,90	86,32	88,48
ALBRK	83,25	83,65	83,48
GARAN	75,92	75,00	76,79
HALKB	73,07	72,87	73,48
ISCTR	78,44	78,01	78,68
SKBNK	88,02	87,46	88,57
TSKB	77,15	76,90	77,89
VAKBN	72,80	72,55	73,46
YKBNK	75,66	75,40	76,79
Avg.	79,13	78,68	**79,74**

and GloVe are limited by the number of words in the training set. Considering that FastText is a model trained at the sub-word level (character-based), every word is learned at the sub-word level, so it can cover much more words than the models at the word level. Also, due to the fact that it is trained at the sub word level, even if a word is misspelled, it can have a very close word representation because of its similarity to the correct word. Thus, this causes the increase of the system performance in terms of classification success. As a result, FastText model provides superior performance in the data sets we presented compared to other models which shows that our model is consistent.

We also employ BerTurk [40] model which is a community-driven cased BERT model for Turkish. Experiment results demonstrate that BerTurk is capable to classify sentiment of users in Twitter data set with 82.25% of accuracy where it provides nearly 2% improvement in classification accuracy compared to the best results of combined models. However, in Mynet Finans, BigPara and KAP data sets, BerTurk performs either almost the same or lower classification success compared to the best performed result of combined models. We consider that it can be related to the term length of the data sets used in this work and pre-trained data sets of BerTurk model. Moreover, to get the results of our experiments in a short time, we conduct BerTurk experiments with TPU, which is produced personally by Google and whose performance is also superior to the GPU.

In [41], Reddy et. al propose the LSTM based hybrid model by combining Word2Vec, GloVe, and FastText models for paraphrase identification in Tamil language. They report that the consolidation of LSTM and FastText model exhibits 73.10% classification accuracy. In our study, proposed model outperforms for all data sets in Turkish. In [42], Arıkan et. al propose an approach that detect and correct the "de/da" clitic errors in Turkish texts. For this purpose, they investigate the impact of different word embedding models such as Word2Vec, GloVe, and FastText. They conclude the study that FastText model presents 64% of accuracy. In our study, the combinations of FastText model with different deep learning models exhibits more successful classification performance. In [43], Doğan and Kaya concentrate on sentiment analysis and text summarization tasks in different social networks such as Twitter, Instagram by employing deep learning techniques. They present the classification results of word embedding models namely, Word2Vec and FastText and their combination results with LSTM model. They report that the combination of Word2Vec and LSTM model exhibits nearly 4% improvement compared to the FastText and LSTM model in the Instagram data set. LSTM+FastText consolidation obtains 83.0% of f-score value while classification performances in our study are 80.23% of accuracy score in Twitter, 86.39% of accuracy value in Mynet Finans, 70.48% of accuracy in Bigpara, and 78.87% of accuracy in KAP data sets. In [44], Bahçevan *et al.* evaluate the neural network performances in Turkish for part-of-speech tagging problem. Unlike aforementioned studies, authors utilize deep neural network (DNN) for the purpose of learning distributed representations that is called word embeddings. LSTM and RNN architectures are employed for comparing the performances of two models for part-of-speech tagging problem. They inform that LSTM 89.0% of accuracy while RNN method exhibits 79.7% of accuracy result. This study also demonstrates that the usage of LSTM model presents remarkable results. In our study, because of the lack of DNN model at the embedding phase, it is hard to compare the performances of these studies. If we use different word embedding model like this work, the consolidation of FastText and LSTM model in Mynet Finans data set exhibits 86.39% of accuracy in our study.

To observe the details of error matrix for each data set Table 11 is presented.

Previous study [45] also concentrates on the direction prediction of large volume bank stocks in BIST 100. Unlike this study, our previous study separately evaluates the impact of deep learning algorithms and a new generation word

TABLE 11. True Positives (TP), False Positives (FP), False Negatives (FN) and True Negatives (TN) values of Actual and Predicted Classes on Error Matrix for each data set.

Dataset	True Positives (TP)	False Positives (FP)	False Negatives (FN)	True Negatives (TN)
Twitter	1322	305	350	1400
Mynet Finans	738	146	147	1124
KAP	450	158	150	631
Bigpara	535	154	153	671

embedding model, Bidirectional Encoder Representations from Transformers (BERT). In this work, we focus on the all combinations of word embedding models and deep learning algorithms. This means that feature selection step of deep learning models is implemented by including word embedding models especially into RNN and LSTM models. In addition to the lack of feature selection steps, the effect of each model (CNN, RNN, LSTM, and BERT) is assessed one by one for the direction prediction of BIST 100 in our previous study. Thus, the main contribution of this study is the investigation of the combinations of deep learning models and word embedding techniques to forecast of the direction of BIST 100. Moreover, previous study only evaluates Twitter data set while three more different platforms are assessed in addition to Twitter in this work. When the experimental results are compared with the previous study over a common data set (Twitter), deep learning techniques exhibit similar classification success.

In [45], training accuracy results are presented in Table 1 while both training and test loss scores are demonstrated Figure 2. When we compare test results of both study at the same epoch size which is 2, deep learning algorithms show either nearly same test accuracy values or approximately between 1% and 8% improvement compared to this study depending on the deep learning method used. In [38], CNN, RNN, and LSTM exhibit 81.23%, 82.95%, 87.44% classification performance, respectively while combinations of FastText and CNN, RNN, LSTM, demonstrate 80.37%, 80.13% and 80.23% classification success in this work, respectively. In this work, after experimenting the other epoch sizes 3, 4, and 5, approximately 2% more decrements are observed for each deep learning algorithm. On the other hand, the same epoch sizes for previous study cause over-fitting. This means stop criterion should be set to 2 for epoch size parameter in previous study. In addition, the performance of BERT is also evaluated in our previous study [38]. However, unlike the other word embedding models, BERT model is designed to pre-train the data set in both layers in two directions and to condition the word in both right and left contexts. Thus, the structure of BERT is wholly different when compared to the conventional word embedding models. Moreover, we cannot compare the performance of BERT with our combined models. In order to make comparisons, it is necessary to combine the BERT model in the same combination way like BERT+deep learning algorithm or to compare it with the other word embedding models. However, if we compare the proposed combined models with the BERT model on the Twitter data set, it offers 84.32% test accuracy when epoch size is set to 2. At the same epoch size, BERT outperforms combined models approximately 4% enhancement.

Considering the Mynet Finans, Bigpara and KAP data sets collected from different platforms, we have found that the results obtained from these data sets are noteworthy in addition to the Twitter data set. Since Mynet Finans is a data set with user comments like Twitter, it is a platform followed by investors with related shares. Therefore, there may be a large number of small investors that will direct their investments by following comments here. With an accuracy of 86.30%, it performs nearly 6% better classification success in the same stocks compared to the Twitter data set. Therefore, it can be a preferred investment tool instead of Twitter. On the other hand, the reason for collecting the Bigpara data set is that investments made with the technical analysis included here can be preferred especially for investors who do not have technical analysis knowledge in determining the direction of the stocks. As a result of our analysis with the data here, we obtained an accuracy of 77.82%. Meanwhile, since the KAP data set contains objective news related to stocks, there are a lot of investors who trade daily just by following this platform. Our test results provide the best 79.74% accuracy value for investors who want to direct their investments by examining this platform.

V. DISCUSSION AND CONCLUSION

In this study, unlike the recent researches on forecasting the stock market direction, we focus on financial sentiment analysis using the Turkish data sets collected from both a social media platform and websites including technical analysis and news to analyze the stock market direction by evaluating high volume stocks in BIST 100. Twitter as a social media platform, KAP as a public disclosure platform, Mynet Finans as an individual comments and actual news platform, and Bigpara as a technical analysis platform on stocks are employed to obtain different text resources using our customized crawler. After gathering and cleaning data sets, the consolidation of word embedding techniques and deep learning approaches are utilized to predict the direction of BIST 100. In addition, stop-word removal, removing of hashtags, removing of URLs, and stemming are used to boost the classification performance of the proposed model at preprocessing step.

To our knowledge, this is the very first attempt in order to forecast the direction of BIST 100 in terms of especially usage of combinations of deep learning models and word embedding techniques. Furthermore, instead of utilizing only BIST 100 data set to estimate movement direction of stock market, the proposed model is focused on 9 high volume stocks which orient the direction of BIST 100. For this purpose, four different platforms are utilized in order to gather text-based news. Analysis of financial news, user comments, the texts of technical analysis through the consolidation of word embedding techniques and deep learning models used in this study will shed light upon investors with a perspective to guide their investment.

REFERENCES

[1] N. Çelik and M. F. Kaya, "Estimation of value at risk by extreme values method and an application on Istanbul stock exchange," *J. Banking Insurance Res.*, vol. 1, no. 5, pp. 19–32, 2010.

[2] V. M. Prieto, S. Matos, M. Alvarez, F. Cacheda, and J. L. Oliveira, "Twitter: A good place to detect health conditions," *PLoS ONE*, vol. 9, no. 1, 2014, Art. no. e86191.

[3] A. Beykikhoshk, O. Arandjelović, D. Phung, and S. Venkatesh, "Overcoming data scarcity of Twitter: Using tweets as bootstrap with application to autism-related topic content analysis," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining*, Paris, France, Aug. 2015, pp. 1354–1361.

[4] A. Bruns, A. Kornstadt, and D. Wichmann, "Web application tests with selenium," *IEEE Softw.*, vol. 26, no. 5, pp. 88–91, Sep. 2009.

[5] S. Loria, "Textblob documentation," Tech. Rep., 2018, pp. 1–73. [Online]. Available: https://textblob.readthedocs.io/en/dev/

[6] Z. H. Kilimci and S. Akyokus, "Deep learning-and word embedding-based heterogeneous classifier ensembles for text classification," *Complexity*, vol. 2018, Oct. 2018, Art. no. 7130146.

[7] H. Gunduz, Y. Yaslan, and Z. Cataltepe, "Intraday prediction of borsa istanbul using convolutional neural networks and feature correlations," *Knowl.-Based Syst.*, vol. 137, pp. 138–148, Dec. 2017.

[8] I. Santos, N. Nedjah, and L. de Macedo Mourelle, "Sentiment analysis using convolutional neural network with fastText embeddings," in *Proc. IEEE Latin Amer. Conf. Comput. Intell. (LA-CCI)*, Arequipa, Peru, Nov. 2017, pp. 1–5.

[9] L. Zhang, S. Wang, and B. Liu, "Deep learning for sentiment analysis: A survey," *Wiley Interdiscipl. Rev., Data Mining Knowl. Discovery*, vol. 8, no. 4, 2018, p. e1253.

[10] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017.

[11] J. L. Elman, "Finding structure in time," *Cognit. Sci.*, vol. 14, no. 2, pp. 179–211, Mar. 1990.

[12] Z. C. Lipton, J. Berkowitz, and C. Elkan, "A critical review of recurrent neural networks for sequence learning," 2015, *arXiv:1506.00019*. [Online]. Available: http://arxiv.org/abs/1506.00019

[13] V. Tunal? and T. T. Bilgin, "PRETO: A high-performance text mining tool for preprocessing turkish texts," in *Proc. 13th Int. Conf. Comput. Syst. Technol.*, Ruse, Bulgaria, 2012, pp. 134–140.

[14] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," 2013, *arXiv:1301.3781*. [Online]. Available: http://arxiv.org/abs/1301.3781

[15] J. Pennington, R. Socher, and C. Manning, "Glove: Global vectors for word representation," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, Doha, Qatar, 2014, pp. 1532–1543.

[16] A. Joulin, E. Grave, P. Bojanowski, and T. Mikolov, "Bag of tricks for efficient text classification," 2016, *arXiv:1607.01759*. [Online]. Available: http://arxiv.org/abs/1607.01759

[17] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," 2018, *arXiv:1810.04805*. [Online]. Available: http://arxiv.org/abs/1810.04805

[18] S. B. Kiliç, S. Paksoy, and T. Genç, "Forecasting the direction of BIST 100 returns with artificial neural network models," *Int. J. Latest Trends Finance Econ. Sci.*, vol. 4, no. 3, pp. 759–765, 2014.

[19] K. Chen, Y. Zhou, and F. Dai, "A LSTM-based method for stock returns prediction: A case study of China stock market," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Santa Clara, CA, USA, Oct. 2015, pp. 2823–2824.

[20] A. Hasan, O. Kalipsiz, and S. Akyokus, "Predicting financial market in big data: Deep learning," in *Proc. Int. Conf. Comput. Sci. Eng. (UBMK)*, Antalya, Turkey, Oct. 2017, pp. 510–515.

[21] N. Pervan and H. Y. Keleş, "Sentiment analysis using a random forest classifier on turkish Web comments," *Commun. Fac. Sci. Univ. Ankara*, vol. 59, no. 2, pp. 69–79, 2017.

[22] S. Tekin and E. Canakoglu, "Analysis of price models in istanbul stock exchange," in *Proc. 27th Signal Process. Commun. Appl. Conf. (SIU)*, Sivas, Turkey, Apr. 2019, pp. 1–4.

[23] S. Tekin and E. Canakoglu, "Prediction of stock returns in istanbul stock exchange using machine learning methods," in *Proc. 26th Signal Process. Commun. Appl. Conf. (SIU)*, Izmir, Turkey, May 2018, pp. 1–4.

[24] H. Gunduz, Z. Cataltepe, and Y. Yaslan, "Stock market direction prediction using deep neural networks," in *Proc. 25th Signal Process. Commun. Appl. Conf. (SIU)*, Antalya, Turkey, May 2017, pp. 1–4.

[25] L. Malandri, F. Xing, C. Orsenigo, C. Vercellis, and E. Cambria, "Public mood–driven asset allocation: The importance of financial sentiment in portfolio management," *Cogn. Comput.*, vol. 10, no. 6, pp. 1167–1176, 2018.

[26] F. Z. Xing, E. Cambria, and R. E. Welsch, "Intelligent asset allocation via market sentiment views," *IEEE Comput. Intell. Mag.*, vol. 13, no. 4, pp. 25–34, Nov. 2018.

[27] A. Picasso, S. Merello, Y. Ma, L. Oneto, and E. Cambria, "Technical analysis and sentiment embeddings for market trend prediction," *Expert Syst. Appl.*, vol. 135, pp. 60–70, Nov. 2019.

[28] X. Li, P. Wu, and W. Wang, "Incorporating stock prices and news sentiments for stock market prediction: A case of hong kong," *Inf. Process. Manage.*, vol. 57, no. 5, Sep. 2020, Art. no. 102212.

[29] Y. Qian, Z. Li, and H. Yuan, "On exploring the impact of users' bullish-bearish tendencies in online community on the stock market," *Inf. Process. Manage.*, vol. 57, no. 5, 2020, Art. no. 102209.

[30] A. C. Turkmen and A. T. Cemgil, "An application of deep learning for trade signal prediction in financial markets," in *Proc. 23nd Signal Process. Commun. Appl. Conf. (SIU)*, Trabzon, Turkey, May 2015, pp. 2521–2524.

[31] T. Mikolov, E. Grave, P. Bojanowski, C. Puhrsch, and A. Joulin, "Advances in pre-training distributed word representations," 2017, *arXiv:1712.09405*. [Online]. Available: http://arxiv.org/abs/1712.09405

[32] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[33] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.

[34] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: A brief review," *Comput. Intell. Neurosci.*, vol. 2018, Feb. 2018, Art. no. 7068349.

[35] D. Kent and F. M. Salem, "Performance of three slim variants of the long short-term memory (LSTM) layer," 2019, *arXiv:1901.00525*. [Online]. Available: http://arxiv.org/abs/1901.00525

[36] M. Sahin. (2019). *HepsiBuradaYorumlar*. Kaggle. Accessed: Jul. 16, 2019. [Online]. Available: https://www.kaggle.com/murats/hepsiburadayorumlar

[37] Z. H. Kilimci and S. Akyokus, "The analysis of text categorization represented with word embeddings using homogeneous classifiers," in *Proc. IEEE Int. Symp. Innov. Intell. Syst. Appl. (INISTA)*, Sofia, Bulgaria, Jul. 2019, pp. 1–6.

[38] Z. H. Kilimci and S. Akyokus, "The evaluation of word embedding models and deep learning algorithms for turkish text classification," in *Proc. 4th Int. Conf. Comput. Sci. Eng. (UBMK)*, Samsun, Turkey, Sep. 2019, pp. 548–553.

[39] Z. H. Kilimci, "Financial sentiment analysis with Deep Ensemble Models (DEMs) for stock market prediction," *J. Fac. Eng. Archit. Gazi Univ.*, vol. 35, no. 2, pp. 635–650, 2020.

[40] S. Schweter, "BERTurk—BERT models for Turkish," Zenodo, 2020, doi: 10.5281/zenodo.3770924.

[41] R. D. Aravinda, K. M. Anand, and K. P. Soman, "LSTM based paraphrase identification using combined word embedding features," *Soft Computing and Signal Processing. Advances in Intelligent Systems and Computing*. Singapore: Springer, 2019, pp. 385–394.

[42] U. Arikan, O. Gungor, and S. Uskudarli, "Detecting clitics related orthographic errors in Turkish," in *Proc.-Natural Lang. Process. Deep Learn. World*, Varna, Bulgaria, Oct. 2019, pp. 71–76.

[43] E. Dogan and B. Kaya, "Deep learning based sentiment analysis and text summarization in social networks," in *Proc. Int. Artif. Intell. Data Process. Symp. (IDAP)*, Malatya, Turkey, Sep. 2019, pp. 1–6.

[44] C. A. Bahcevan, E. Kutlu, and T. Yildiz, "Deep neural network architecture for Part-of-Speech tagging for turkish language," in *Proc. 3rd Int. Conf. Comput. Sci. Eng. (UBMK)*, Sep. 2018, pp. 235–238.

[45] D. Othan, Z. H. Kilimci, and M. Uysal, "Financial sentiment analysis for predicting direction of stocks using bidirectional encoder representations from transformers (BERT) and deep learning models," in *Proc. Int. Conf. Innov. Intell. Technol.*, Istanbul, Turkey, vol. 2019, pp. 30–35.

ZEYNEP HİLAL KILIMCI was born in Erzurum, Turkey, in 1985. She graduated from the Computer Engineering Department, Doğuş University, in 2008, the M.Sc. degree from the Computer Engineering Department, Doğuş University, in 2013, and the Ph.D. degree from the Computer Engineering Department, Kocaeli University, in 2018. From 2009 to 2011, she worked as a Software Engineer with the Data Mining & CRM Department, DenizBank. In 2011, she started to work as a Research Assistant with the Computer Engineering Department, Doğuş University, where she also worked as an Assistant Professor, from 2018 to 2020. She is currently working as an Assistant Professor with Kocaeli University. Her research fields include text mining, data mining, machine learning, ensemble learning, deep learning, reinforcement learning, and artificial intelligence.

RAMAZAN DUVAR was born in Isparta, Turkey, in 1988. He received the B.Sc., M.Sc., and Ph.D. degrees in electronics and telecommunication engineering from the University of Kocaeli, Kocaeli, Turkey, in 2010, 2012, and 2019, respectively.

From 2012 to 2019, he was a Research Assistant with the Electronics and Telecommunication Engineering Department, Kocaeli University. Since 2020, he has been an Assistant Professor with the Aircraft Electrical and Electronics Department, Kocaeli University. His research interests include digital signal, image/video processing, embedded systems, machine learning, and deep learning.