Heliyon 10 (2024) e23784

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Sentiment analysis of financial Twitter posts on Twitter with the machine learning classifiers

Handan Cam [a], Alper Veli Cam [b], Ugur Demirel [c,**], Sana Ahmed [d,*]

[a] *Department of Management Information Systems, Faculty of Economic and Administrative Science, Gumushane University, 29000, Gumushane, Turkey*

[b] *Department of Health Care Management, Faculty of Health Sciences, Gumushane University, 29000, Gumushane, Turkey*

[c] *Irfan Can Kose Vocational School, Gumushane University, 29000, Gumushane, Turkey*

[d] *Henley Business School, University of Reading, Reading, RG6 6AH, UK*

ARTICLE INFO

Keywords:
Sentiment analysis
Natural language processing
Machine learning
Stock market
Twitter

ABSTRACT

This paper presents a sentiment analysis combining the lexicon-based and machine learning (ML)-based approaches in Turkish to investigate the public mood for the prediction of stock market behavior in BIST30, Borsa Istanbul. Our main motivation behind this study is to apply sentiment analysis to financial-related tweets in Turkish. We import 17189 tweets posted as "#Borsaistanbul, #Bist, #Bist30, #Bist100″ on Twitter between November 7, 2022, and November 15, 2022, via a MAXQDA 2020, a qualitative data analysis program. For the lexicon-based side, we use a multilingual sentiment offered by the Orange program to label the polarities of the 17189 samples as positive, negative, and neutral labels. Neutral labels are discarded for the machine learning experiments. For the machine learning side, we select 9076 data as positive and negative to implement the classification problem with six different supervised machine learning classifiers conducted in Python 3.6 with the sklearn library. In experiments, 80 % of the selected data is used for the training phase and the rest is used for the testing and validation phase. Results of the experiments show that the Support Vector Machine and Multilayer Perceptron classifier perform better than other classifiers with 0.89 and 0.88 accuracy and AUC values of 0.8729 and 0.8647 respectively. Other classifiers obtain approximately a 78,5 % accuracy rate. It is possible to increase sentiment analysis accuracy with parameter optimization on a larger, cleaner, and more balanced dataset by changing the pre-processing steps. This work can be expanded in the future to develop better sentiment analysis using deep learning approaches.

1. Introduction

The increasing use of social media platforms has made it possible for people to express their opinions on social blogs such as Facebook and Twitter [1]. Thus, understanding the emotions in these opinions has become an important issue for researchers, but the increasing volumes of these opinions make it impossible for a human process which led to a need for automated processing [2]. For this purpose, natural language processing (NL) and artificial intelligence techniques have been used to conduct automatic sentiment

* Corresponding author.
** Corresponding author.
E-mail addresses: hcam@gumushane.edu.tr (H. Cam), alpercam@gumushane.edu.tr (A.V. Cam), udemirel@gumushane.edu.tr (U. Demirel), sanaazlan4545@gmail.com (S. Ahmed).

https://doi.org/10.1016/j.heliyon.2023.e23784
Received 25 May 2023; Received in revised form 5 December 2023; Accepted 13 December 2023
Available online 17 December 2023
2405-8440/

analysis to determine the emotion specified in any given source.

Sentiment analysis, also called Opinion Mining, is one of the most popular techniques for the computational study of opinions and sentiments of people's thoughts, perceptions, and feelings expressed in the text [3,4]. Even though there are studies on stock market predictions [5] questioning whether the stock market can be predicted with the help of sentiment analysis or not, most works in this field have been carried out on the evaluation of consumer products and movie reviews [6]. Thus, the rationale behind our decision to employ sentiment analysis on financial-related tweet data in Turkish stems from the limited existing research in this particular domain.

The study aims to conduct a sentiment analysis combining the lexicon-based and machine learning (ML)-based approaches in Turkish to investigate the public mood for the prediction of stock market behavior in BIST30, Borsa Istanbul. This paper presents a sentiment analysis model for Turkish texts by combining lexicon-based methods and machine-learning classifiers. The dataset is obtained from instant tweets with MAXQDA 2020 between November 7, 2022, and November 15, 2022, to be utilized as input for machine learning models. MAXQDA is a computer-assisted data analysis software application that has been developed by VERBI Software, a company headquartered in Berlin, Germany [7]. We use Orange, open-source machine learning and data visualization software [8] to extract sentiments from tweets and assign appropriate labels. Orange offers multilingual sentiment including sentiment scores of the words and preprocessing module for the pretreatment processes of Turkish texts for lexicon-based method. After the pretreatment processes of Turkish texts, machine learning models have been applied by using lexicon-based features of texts to conduct sentiment analysis for unlabeled tweet data on company shares between 21.11.2022 and 25.11.2022. The main contributions of this study are as follows:

1 A new dataset of financial tweets in Turkish has been obtained
2 Preprocessing algorithms in Turkish have been implemented to prepare sentences to be sent into the sentiment analysis algorithm
3 Six machine-learning algorithms have been implemented for sentiment analysis.
4 These algorithms have been tested and compared on this dataset

The remainder of this paper is organized as follows. In Section 2, we briefly discuss related work on SA. Section 3, presents the methodology. In section 4, we consider the experimental setup, such as datasets and evaluation metrics. In section 5, we evaluate the results and in section 6, we discuss limitations. Finally, in Section 7, we discuss the conclusion and future work.

2. Related work

The use of machine learning techniques for real-time predictions and the adoption of these techniques, including Naive Bayes (NB), Logistic Regression (LR), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Decision Trees (DTs), and Multilayer Perceptron (MLP), have been prominent in the literature. Numerous studies [9–12] have provided evidence of the efficacy of these methodologies in diverse forecasting scenarios. Numerous research have been conducted utilizing the aforementioned methodologies for sentiment analysis as well.

Sentiment analysis research is reviewed under three headings: machine-learning methods, lexicon-based methods, and hybrid methods [13]. Machine learning (ML) methods implement well-known ML algorithms and use linguistic features. Lexicon-based methods (LB), divided into dictionary-based methods and corpus-based methods using statistical or semantic methods to find sentiment polarity, rely on sentiment dictionaries. Hybrid methods combining the two methods are widely used in sentiment dictionaries and play an important role in most methods [14]. Many studies have been performed using the above-mentioned methods for sentiment analysis in various languages, especially in English [15].'s work with machine learning algorithms can be cited as examples of pioneering studies using machine learning methods for sentiment analysis. Studies, such as [4,15–18] conducted various machine learning (ML) and deep learning (DL) methods for sentiment analysis and sentiment classification on different sentiment analysis datasets in English; whereas studies [19,20] tested machine learning algorithms with the help of a lexicon which improved the accuracies of the classifiers to determine sentiment in Arabic datasets obtained from different platforms. Besides Arabic, studies [21,22] applied machine learning algorithms to user reviews in Chinese datasets for emotion classification. In both studies, it was proven that the NB classification algorithm achieved better accuracy than the SVM algorithm. As for the machine learning approach in Turkish, studies such as [23–28] used ML and DL methods for sentiment analysis on Turkish datasets obtained from different platforms. However, recent research has focused on utilizing Twitter as a primary source for sentiment analysis, since it serves as a comprehensive store of individuals' emotional datasets.

Twitter users exhibit prompt responses to matters such as democratic events, movie box-office performance, stock market trends, product and service quality, etc. furnishing noteworthy insights and public sentiments. The analysis of these comments enables the assessment of public responses. For example, the studies [29–35] conducted sentiment analyses on tweet data related to partisan news, cybercriminals and politicians, West Java Governor Election, US Airline Companies, IMDB datasets, electronic product reviews, and movie and restaurant review data. Recently, several studies, such as [36–42], have focused on exploring the performance of ML and DL classifiers for sentiment analysis of COVID-19 tweets. Studies [1,43–46] might likewise be considered examples of Twitter data analysis. Although sentiment analysis has been extensively studied in several domains, there is a dearth of research on its application specifically to sentiment analysis of financial-related tweets for stock market forecasting. The objective of our research is to address this disparity by concentrating on the sentiment that is articulated in a manner relevant to a certain domain, thereby providing insights into its influence on market dynamics. As for the sentiment analysis with ML and DL methods on Turkish tweets, studies such as [26,47–55] conducted experiments on a Twitter dataset with different classifiers to detect emotions [48]. presented SA analysis using NB, RF, and SVM on Twitter data with an achievement of 90 % F1 score [49]. obtained Turkish data set from Twitter and applied NB, multinominal Naïve Bayes (MNB), SVM, and k-NN algorithms for classification by using BoW and n-gram models as feature extraction [50]. performed sentiment analysis by conducting experiments on a Twitter dataset with different classifiers. They produced the best results using MLP and SVM with more than 80 % accuracy. Our experimental findings indicate that the Support Vector Machine and Multilayer Perceptron classifier exhibit superior performance compared to other classifiers, achieving an accuracy of 0.89 and 0.88 and AUC values of 0.8729 and 0.8647, respectively. Other classifiers achieve an accuracy rate of roughly 78.5 %.

The utilization of lexicon-based approaches alongside machine learning methods has been observed in sentiment analysis inside the literature. The lexicon-based technique is based on a lexicon created from a corpus or dictionary, where words expressing the identified emotional states and their synonyms are looked up, and an estimated sentiment score is calculated [56]. created SentiWordNet to assign three polarity scores as positive, negative, and neutral to each English WordNet synset [57]. obtained sentiment scores from SentiWordNet to use in their opinion-mining application by creating movie scores from blog pages [58]. proposed a four-step approach as sentiment word detection, sentiment shifter detection, sentiment score handling, and aggregated score calculation for aspect-sentiment matching on customer reviews of products [59]. performed hybrid lexicon-based sentiment analysis on tweet data with their polarities gained from Sentiment140.com. As for the lexicon-based approach in Turkish [60], developed wordnet for Turkish within the scope of the Balkanet project by translating synset in WordNet [61]. presented a system for sentiment analysis by translating the SentiStrength lexicon into Turkish [62]. implemented the LB method and n-gram methods on Twitter data and found that the LB method had better performance compared to the n-gram method [63]. developed a Turkish polarity lexicon known as the SentiTurkNet assigning polarities to each synset in Turkish WordNet. Later on [64], proposed and evaluated a Turkish sentiment analysis system covering different levels of SA along with some linguistic issues in Turkish movie reviews [65]. determined the sentiment polarity of Turkish and English tweets with the lexicon-based sentiment analysis applied to Turkish and English Twitter messages about Syrian refugees and revealed the differences in polarity in both languages [66]. introduced a lexicon-based SA for Turkish using different Twitter datasets and concluded that their proposed system obtained better achievement than other previously developed Turkish lexicon-based SA systems. Apart from the lexicon-based methods mentioned above, we utilized a multilingual sentiment method provided by the Orange program [67] for several languages including the Turkish language to represent the sentiment score of a tweet as positive, negative, or neutral.

The hybrid approach combines lexicon-based methods with machine learning techniques [68]. determined the sentiment using SentiWordNet Lexical resource for sentiment analysis of film reviews with an accuracy of 69.35 % and then applied one of the machine learning methods like the SVM classifier [69]. proposed a hybrid method combining NB and GA (genetic algorithm) on movie reviews for sentiment analysis and observed that NB and GA performed worse than the hybrid method when they were applied separately [2]. presented a hybrid method for SA on movie reviews by using machine learning techniques, semantic rules, fuzzy sets, and a sentiment lexicon as well [70]. used machine learning algorithms and lexicon-based approaches together on the product comments and evaluations with 73 % success without the need for human involvement [71]. performed a sentiment analysis on three different data sets with a hybrid approach using lexicon-based and machine-learning approaches together with an accuracy of 7 % on average.

The current body of scholarship about stock market forecasting frequently fails to adequately consider the possible insights that might be derived from sentiment analysis on social media. The objective of our research is to address this gap in knowledge by examining the incorporation of sentiment data, specifically sourced from Twitter, into predictive models. To the best of our knowledge, there exists a scarcity of research investigating the potential of utilizing sentiment analysis as a means to forecast stock market behavior. The study [5] examines the potential correlation between collective mood states derived from Twitter feeds by employing mood-monitoring technologies to examine and evaluate the daily Twitter feeds. Another study [72], investigates the correlation between Turkish tweets and the Turkish stock market index by employing data mining techniques on a dataset sourced from Twitter in Türkiye; whereas our research endeavor represents the initial attempt to construct and assess a hybrid framework that examines and categorizes the emotional expressions conveyed in tweets related to the stock market. This analysis is conducted through the utilization of a lexicon-based approach, specifically employing the multilingual sentiment tool, and subsequently employing cutting-edge machine-learning approaches to classify the tweets in Turkish.

3. Methodology

A conceptual model of the procedures to be followed in performing the proposed sentiment analysis on Twitter data is shown in Fig. 1.

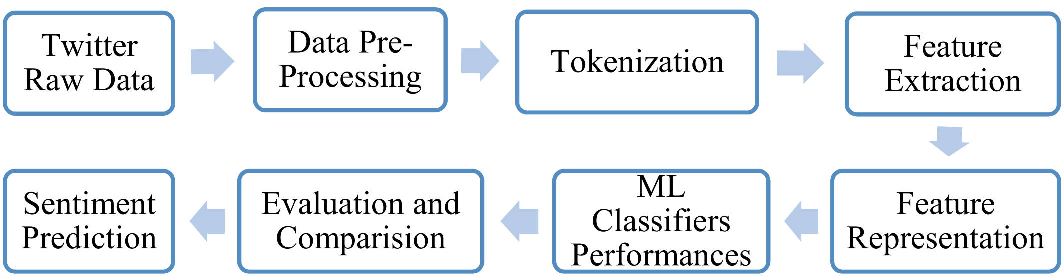

Fig. 1. Sentiment analysis conceptual diagram.

In this study, the MAXQDA 2020 program has been used to import 17189 tweets posted as "#Borsaistanbul, #Bist, #Bist30, #Bist100″ on Twitter between November 7, 2022, and November 15, 2022, to be utilized as input for machine learning models. To import Twitter data, from the import section in the program, the Twitter Data button needs to be clicked and then a dialog window demonstrated in Fig. 2 appears to get a Twitter account linked to MAXQDA. Once Twitter is linked using login data of the Twitter account, it must be authorized by the application demonstrated in Fig. 2 to allow MAXQDA to receive data from the Twitter account [73].

The Orange Program, an open-source machine learning, and data mining program, is employed to analyze the data from Twitter. The analysis design is given in Fig. 3. After the data is seen in the "Corpus Viewer" stage, the data is cleaned with preprocessed text to make it better quality. This stage is an important step in any application related to text mining or data mining; there are four processes: transformation, tokenization, normalization, and filtering. Transformation is the process of transforming input data by turning text to lowercase, removing accents, parsing HTML, and removing URLs. Tokenization is the method of splitting text into a series of meaningful parts called tokens. Normalization is used to apply stemming and lemmatization to words. Filtering is often used to filter out unnecessary words that appear in a text document. After having done the preprocess text step, word cloud, and sentiment analysis steps were carried out with the cleaned data [74]. Sentiment analysis is used for sentiment prediction from a corpus loaded with Twitter data saved in Excel file format. As seen in Fig. 3, the corpus is connected to the sentiment analysis to represent the sentiment score of a tweet as positive, negative, or neutral with a multilingual sentiment method for several languages including the Turkish language [67]. After the preprocessing step is done, the text data can be converted into separate texts and displayed as the word cloud given in Fig. 4. When Fig. 4 is examined, it is seen that the top words in tweets are "Bist100, Bist, Borsa, Hisse, Borsaistanbul, Bist30". The reason behind this is because of the hashtagged words we used when importing the tweet posts. However, it is noteworthy that apart from these hashtagged words, there are other important words giving insights about stock market suggestions in tweet posts such as "borsa, hisse, endeks, yatırım, tavsiye, devam, destek, direnc, bereketli, güzel" ("stock market, share, index, investment, advice, continuation, support, resistance, fertile, beautiful" in English respectively). Here it is seen that; the sample dataset obtained is good enough to develop a model for sentiment analysis in machine learning methods.

The emotions expressed according to the distinction of the tweets as positive, negative, and neutral are presented in Fig. 5. 34,81 % of these tweets are positive, 47,20 % are neutral, and 17,99 % are negative. It can be said that the tweets posted with the hashtags "#Borsaistanbul, #Bist, #Bist30, #Bist100 in the considered time are more positive. However, these feelings will inevitably change as the agenda and time change. Tweets containing the above-mentioned hashtags were collected throughout the time frame of November 7, 2022, to November 15, 2022, utilizing the MAXQDA 2020 software. The multilingual sentiment approach given in the Orange Data Mining application was utilized to categorize the sentiments of tweets into positive, negative, and neutral polarities. In this study, a total of 17,189 tweets were analyzed. From this dataset, 9076 samples were chosen to serve as negative and positive examples for training and testing a machine learning model. The dataset was subjected to a train-test split using the 80-20 rule. The training dataset comprised 80 % of the total data, while the remaining 20 % was allocated for testing (10 %) and validation (10 %) purposes. The tweets from the test dataset were not utilized in any part of the training or validation processes to mitigate the risk of potential model overfitting.

Six supervised machine learning models have been applied by using lexicon-based features of texts to conduct sentiment analysis for unlabeled tweet data on company shares between 21.11.2022 and 25.11.2022. These machine learning classifiers explained below are Naive Bayes (NB), Logistic Regression (LR), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Decision Trees (DTs), and Multilayer Perceptron (MLP).

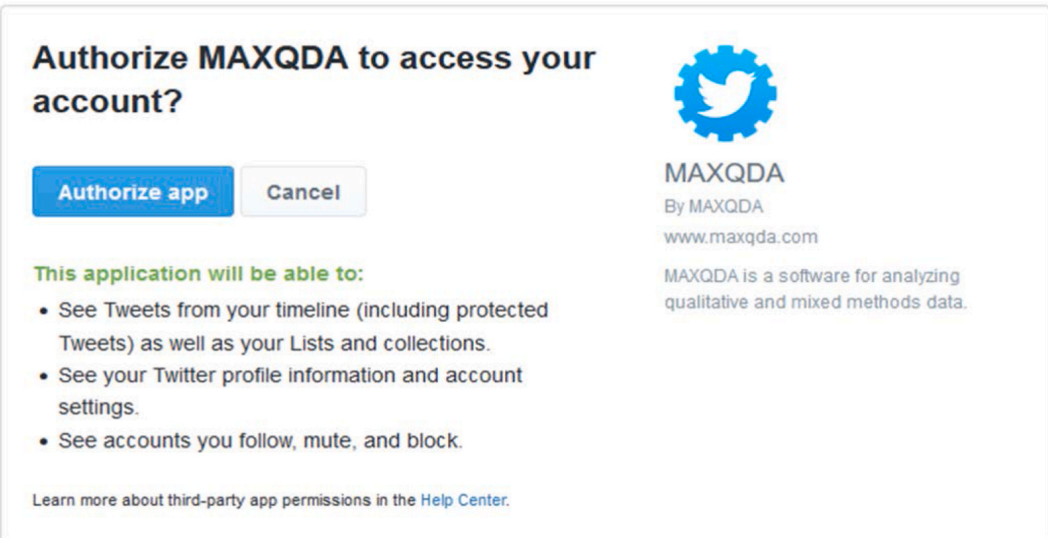

Fig. 2. Importing Twitter data using MAXQDA 2020.

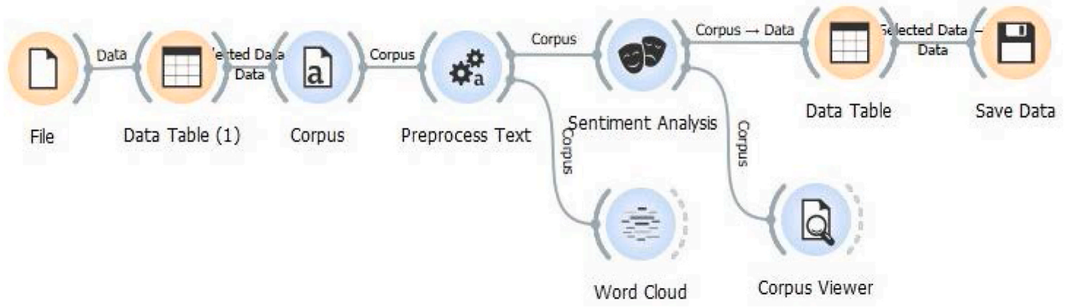

Fig. 3. Lexicon-based technique analysis design.

Fig. 4. Word cloud.

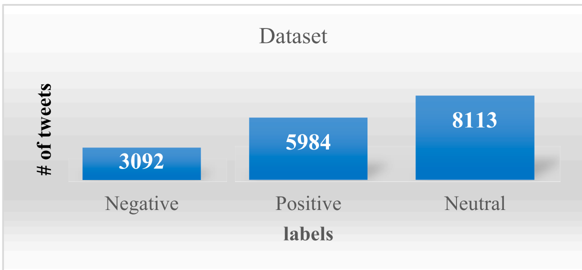

Fig. 5. Dataset characteristics.

3.1. *Naive Bayes (NB) classifier*

The supervised machine learning method Naive Bayes makes assumptions that features are independent and equally important. Even though the features are not independent, the Naive Bayes algorithm is still applicable to many machine learning problems. Naive Bayes is a simple algorithm based on probability theory compared to other complex machine learning algorithms [51]. Bayes rule is defined as in equation (1);

$$\mathrm{P}(\mathrm{c}|\mathrm{t}) = \frac{\mathrm{P}(\mathrm{c})\mathrm{P}(\mathrm{t}|\mathrm{c})}{\mathrm{p}(\mathrm{t})} \tag{1}$$

where c is a specific class (either positive or negative) and t is text classified. P(c) and P(t) are the prior probabilities of c (class) and t (text). P(t|c) is the appearance probability of text in a specified class [1].

3.2. Logistic regression (LR) classifier

Logistic regression estimates an outcome's probability using just two values. With values that can be expressed in a binary manner, such as yes/no and true/false, linear regression is ineffective as it generates a logistic curve with values between 0 and 1. In contrast to linear regression, which uses a probability curve to build its model, logistic regression uses the natural logarithm of the target variable's odds. The formula of logistic regression can be explained in equation (2) [75];

$$\mathrm{Logit}(\mathrm{p}) = \log\left(\frac{\mathrm{p}}{1-\mathrm{p}}\right) \tag{2}$$

3.3. Support vector machine (SVM) classifier

Support vector machine (SVM) is another supervised machine learning method used for classification problems. SVM constructs a line at the greatest distance between points of both classes (positive and negative in our example) on a plane. The widest frontier known as the margin hyper-plane between positive and negative classes has been found by optimizing with quadratic constraint [19,29]. For this purpose, let $\mathrm{y_I}$ equals $\{1,-1\}$ if document $\overrightarrow{d_i}$ is in $\{1,-1\}$ class. The solution can be formulated as:

$$\overrightarrow{w} = \sum_{i=1}^{n} a_i y_i \overrightarrow{d}_i \; a_i \geq 0 \tag{3}$$

Where a_i is obtained by solving a dual optimization problem. The only document vectors that contribute to $\overrightarrow{w}$ are those in which a_i is greater than zero called support vectors [15,21].

3.4. K-nearest neighbor (KNN) classifier

The k-nearest neighbor method classifies newly incoming data using previously categorized data on the basis that the closest data points all belong to the same class. Test samples relate to the unclassified data, whereas learning samples refer to the data that have already been categorized. After determining the distance between the test sample and the learning sample, the KNN technique selects the k-learning sample that is closest to the test sample. Additionally, the class to which the majority of the chosen k samples belonged (the test sample class) is determined [76].

3.5. Decision trees (DTs) classifier

The decision tree is a hierarchical supervised machine learning model to categorize data by performing learning inductively from known data classes. It displays a tree-like structure with leaf nodes serving as the class labels and internal nodes indicating the test conditions. The data space is divided recursively until the leaf nodes have a certain number of records that are utilized for categorization [14,30].

3.6. Multilayer perceptual (MLP) classifier

The multilayer perceptual neural network (MLP) model is the most commonly used artificial neural network model created by connecting neurons in one layer to neurons in the following layer. It consists of three different layers: the input layer, the hidden layer, and the output layer. The input layer is the layer from which the data is read. Since each neuron represents a feature, it contains as many neurons as the number of features. The output layer is the layer where the classes are determined [50].

We compare the findings by examining their accuracy and f1scores following the pre-processing and modeling phases. To determine these numbers, we employ a confusion matrix. To determine the precision, recall, accuracy, and f1 score, a confusion matrix is used. These metric values are obtained from the training subset. Maximum True Negative and True Positive values are desired. True Negative denotes situations in which both the actual and expected data are negative (0). True Positive refers to situations where both the actual and expected data are positive (1). To determine whether there is an overload on one feeling, it is also crucial to evaluate the False Positive and False Negative numbers. In the equations below, True Negative (TN), True Positive (TP), False Negative (FN), and False Positive (FP) are used.

Accuracy is the ratio of the number of correctly classified samples (TP + TN) to the total number of samples (TP + TN + FP + FN).

$$Accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN)} \tag{4}$$

Precision is the ratio of the number of True Positive (TP) samples predicted to the total number of samples predicted (TP + FP).

$$Precision = \frac{(TP)}{(TP + FP)} \tag{5}$$

Recall is the ratio of the number of correctly classified positive samples (TP) to the total number of positive samples (TP + FN).

$$Recall = \frac{(TP)}{(TP + FN)} \tag{6}$$

F1 Score is the harmonic mean of precision and sensitivity measures which are not sufficient alone to draw a meaningful comparison result. Thus, evaluating both criteria together gives more accurate results.

$$F1_{score} = 2x\left(\frac{Precision\ x\ Recall}{Precision + Recall}\right) \tag{7}$$

The area under the ROC curve (AUC) is another criterion used to evaluate the performance of classification algorithms. AUC takes a value in the range of 0–1, and high values indicate that the classification algorithm has a higher estimation performance [29].

The Precision-Recall curve is a visual depiction that illustrates the relationship between precision and recall at varying probability thresholds. The Area Under the Precision-Recall Curve (AUPR) is a numerical metric that provides a concise representation of the model's overall performance across a range of thresholds. The area under the precision-recall curve (AUPR) serves as a measure of the model's ability to effectively distinguish between the positive and negative classes. A greater AUPR value is indicative of superior performance [77].

4. Experimental results

In this study, experiments were performed on a laptop with Intel Core i7-8550U CPU 1.80 GHz, 8 GB RAM, and NVIDIA GeForce MX150 with 2 GB RAM. The applications were implemented in Python 3.6 with the sklearn library. 9076 samples were selected for positive and negative labels for bi-class classification. The neutral-labeled tweets were discarded for the experiments as in studies [31, 70]. ML models in our study are designed to detect and evaluate the attitudes that are not neutral because they often lack unambiguous positive or negative. Discarding these data points simplifies the process by eliminating potentially ambiguous or uninformative information. For machine learning sentiment analysis, 5984 positively and 3092 negatively labeled tweets provided by the multilingual sentiment method in the Orange program were used in experiments. 80 % of our data for the training phase and remaining of the data set was used in the testing (10 %) and validation (10 %) phases. Having done experiments, new financial tweets between November 21, 2022, and November 25, 2022, about the firms listed in Bist30, were given to machine learning algorithms to predict the sentiments of the tweets. Each ML model was evaluated by considering precision, recall, F1-score, accuracy, and AUC value. The parameter tuning for each model was performed by trying different parameters that could improve the performance of the models.

4.1. Naive Bayes (NB) classifier

NB constructs a line at the greatest distance between points of both classes (positive and negative in our example) on a plane. The proportion of all characteristics with the highest variance that are added to variances for calculation stability is set to default as 1e-9. In Fig. 6, the confusion matrix and ROC curve obtained with the Naive Bayes classifier show our results. Accuracy, AUC and AUPR values are 68 %, 0.7297, and 0.8794 respectively. Precision, recall, and F1-score are 52 %, 88 %, and 65 respectively for negative classes and 90 %, 68 %, and 71 % for positive classes as presented in Table 1.

4.2. Logistic regression (LR) classifier

We vectorized our sentences using tf-idf. Then, we ran a hyperparameter search in a grid. For parameter tuning, regularized logistic regressing was implemented by using the 'lbfg' as solver. The maximum number of iterations taken for the solver to converge was set to the default number of 100. In Fig. 7, the confusion matrix and ROC curve obtained with Logistic Regression show our results. Accuracy,

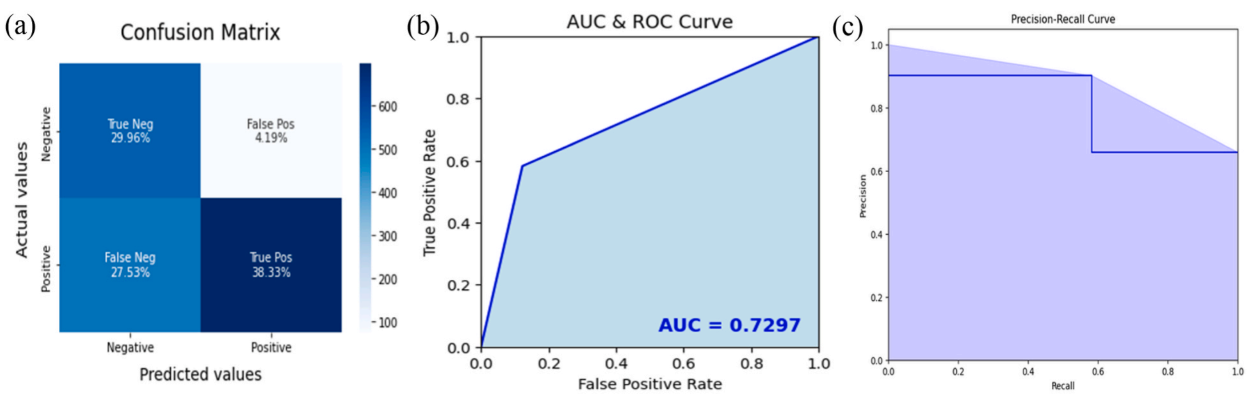

Fig. 6. Performance Metrics for Naive Bayes (NB) classifier. (a) Confusion matrix, (b) AUC & ROC curve and (c) Precision-Recall curve.

AUC and AUPR values are 88 %, 0.8506, and 0.9316 respectively. Precision, recall, and F1-score are 83 %, 81 %, and 82 % respectively for negative classes and 90 %, 91 %, and 91 % for positive classes as seen in Table 2.

4.3. Support vector machine (SVM) classifier

Having done the preprocessing and vectorization steps, the SVM model was implemented based on libsvm with the kernel type of linear. In Fig. 8, the confusion matrix and ROC curve obtained with SVM show our results. Accuracy, AUC and AUPR values are 89 %, 0.8729, and 0.9415 respectively. Precision, recall, and F1-score are 85 %, 82 %, and 84 % respectively for negative classes and 91 %, 93 %, and 92 % for positive classes as presented in Table 3.

4.4. K-nearest neighbor (KNN) classifier

For the KNN classifier, the number of neighbors to use was set to 7 and the weight function used in prediction was the default value set as uniform. The algorithm used auto parameters to compute the nearest neighbors. The Metric to use for distance computation was Minkowski. In Fig. 9, the confusion matrix and ROC curve obtained with KNN show our results. Accuracy, AUC and AUPR values are 78 %, 0.6857, and 0.8616 respectively. Precision, recall, and F1-score are 41 %, 36 %, and 38 % respectively for negative classes and 85 %, 88 %, and 86 % for positive classes as presented in Table 4.

4.5. Decision trees (DTs) classifier

In DTs classifier, gini criteria was used to measure the quality of a split, and the "best" strategy was to choose the best split for parameter tuning. In Fig. 10, the confusion matrix and ROC curve obtained with the DTs classifier show our results. Accuracy, AUC and AUPR values are 81 %, 0.8070, and 0.9073 respectively. Precision, recall, and F1-score are 51 %, 59 %, and 55 % respectively for negative classes and 90 %, 87 %, and 88 % for positive classes as presented in Table 5.

4.6. Multilayer perceptual (MLP) classifier

MLP classifier has many parameters such as hidden layer sizes, activation, solver, etc. The Relu activation function was used for the hidden layer. Hidden layer sizes were 100. For weight optimization, a stochastic gradient-based optimizer known as adam solver was chosen with a number of 200 iterations. In Fig. 11, the confusion matrix and ROC curve obtained with MLP show our results. Accuracy, AUC and AUPR values are 88 %, 0.8647, and 0.9383 respectively. Precision, recall, and F1-score are 84 %, 81 %, and 83 % respectively for negative classes and 90 %, 92 %, and 91 % for positive classes as presented in Table 6.

After the completion of all six different experiments, a comparison of the results was presented in Table 7. When the study's findings were analyzed, the Support Vector Machine and Multilayer Perceptron classifier produced the best results with 0.89 and 0.88 accuracy and the AUC values of 0.8729 and 0.8647 respectively.

Table 8 shows the number of tweets posted between 21.11.2022 and 25.11.2022 about company shares with weekly opening and closing prices. In general, the most mentioned company shares were the ones that had fewer fluctuations and the least mentioned company shares were the ones that had more fluctuations in the prices during the week as seen in Table 8. A plausible rationale for the observed relationship is that greater public attention is directed towards the company shares that are most frequently mentioned. The heightened level of scrutiny could potentially foster a more consistent perception of the market, wherein investors may exhibit reduced impulsive responses to transient developments or fluctuations. Conversely, the least mentioned company shares may encounter challenges stemming from a dearth of information or limited visibility within the public sphere. Investors may exhibit a heightened dependence on speculative activities, resulting in increased levels of price volatility as a consequence of the prevailing ambiguity and absence of consensus over the prospects of the company. The most mentioned company shares used in this study are Sasa Polyester Sanayi AS (SASA), Hektas Ticaret AS (HEKTS), Eregli Demir ve Celik Fabrikalari TAS (EREGL), and Aselsan Elektronik Sanayi ve Ticaret AS (ASELS).

From 21.11.2022 to 25.11.202, BIST30 finished the week at 5265,19 with a 7,05 % increase with the lowest level of 4820,36 and the highest level of 5358,31. As of the closing day of the week, BIST30 finished the day at a 5265,19 level with a 0.22 % increase in the last trading day. The index saw the lowest level of 5185,19 and the highest level of 5287,00 [78].

For prediction purposes, only the most mentioned company shares on Twitter were applied in all six machine learning algorithms. As shown in Table 8, most of the tweets have been posted about SASA, HEKTS, EREGL, and ASELS with 10000, 7377, 4223, and 2588 tweets respectively. There might have been more than 10,000 tweets on SASA, however, MAXQDA could import up to 10,000 tweets at one time. In Table 9, both positive and negative class predictions have been demonstrated to give an insight into the price index direction as shown in Fig. 12.

Table 1
Experiment results for Naive Bayes (NB) classifier.

	Precision	Recall	F1-score	Accuracy	AUC	AUPR
Negative	52 %	88 %	65 %	68 %	0.7297	0.8794
Positive	90 %	68 %	71 %			

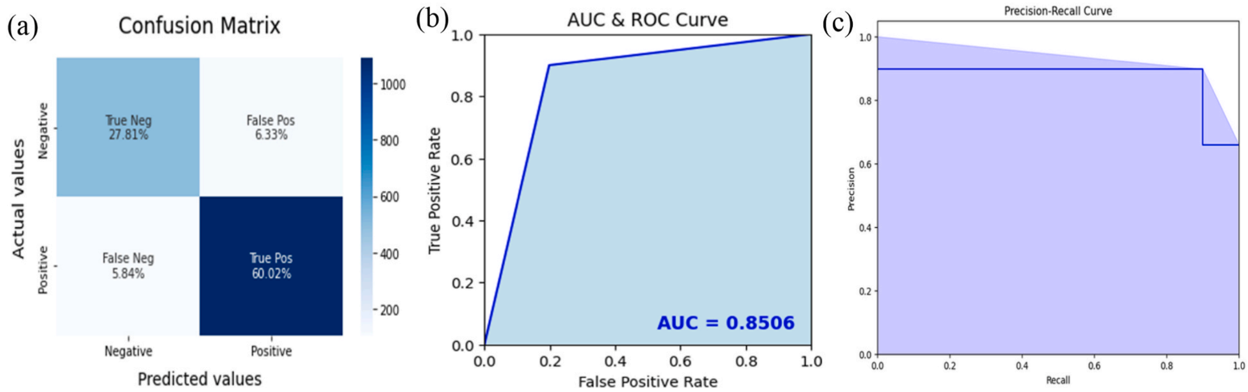

Fig. 7. Performance Metrics for Logistic Regression (LR) classifier. (a) Confusion matrix, (b) AUC & ROC curve and (c) Precision-Recall curve.

Table 2
Experiment results for Logistic Regression (LR) classifier.

	Precision	Recall	F1-score	Accuracy	AUC	AUPR
Negative	83 %	81 %	82 %	88 %	0.8506	0.9316
Positive	90 %	91 %	91 %			

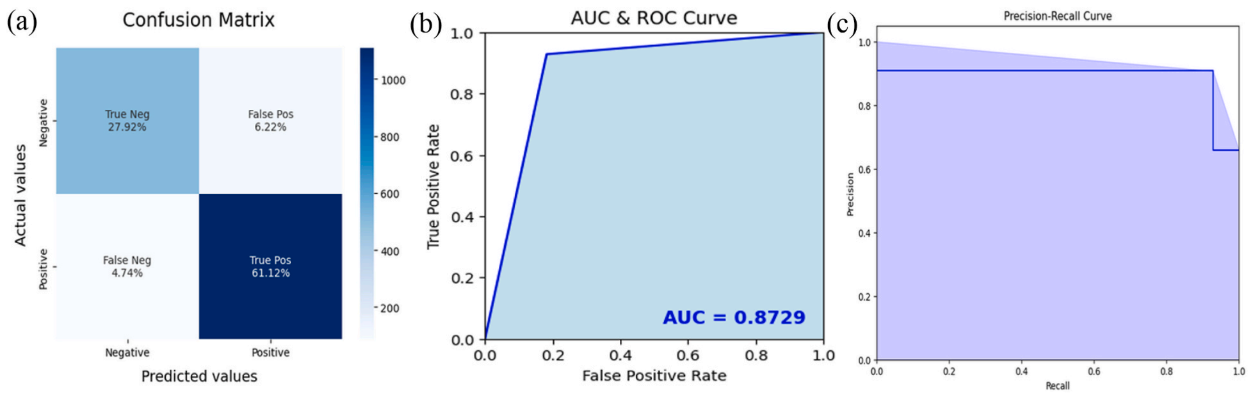

Fig. 8. Performance Metrics for Support Vector Machine (SVM). (a) Confusion matrix, (b) AUC & ROC curve and (c) Precision-Recall curve.

Table 3
Experiment results for Support Vector Machine (SVM) classifier.

	Precision	Recall	F1-score	Accuracy	AUC	AUPR
Negative	85 %	82 %	84 %	89 %	0.8729	0.9415
Positive	91 %	93 %	92 %			

It has been observed that the increases were in the majority across the BIST30 index. In Fig. 12, four companies that contributed positively and negatively are presented. ASELS and EREGL have contributed to the index positively with an increase of 20.26 % and 12.26 % respectively. HEKTS and SASA have contributed to the index negatively with a decrease of 6.47 % and 4.19 % respectively.

5. Discussion

On the social media platform Twitter, individuals have the ability to promptly express their opinions on many topics. These comments provide us with information and insight into the public's reaction regarding the subject. Sentiment Analysis is used to explore the feelings of individuals about their experiences through texts and to identify relevant issues. In addition, since the keywords related to the subject are presented, the relationship of the emotion with certain objects of attraction can be noticed. The sentiment analysis in this study can be used to find information about the price directions of the company shares since people post tweets when

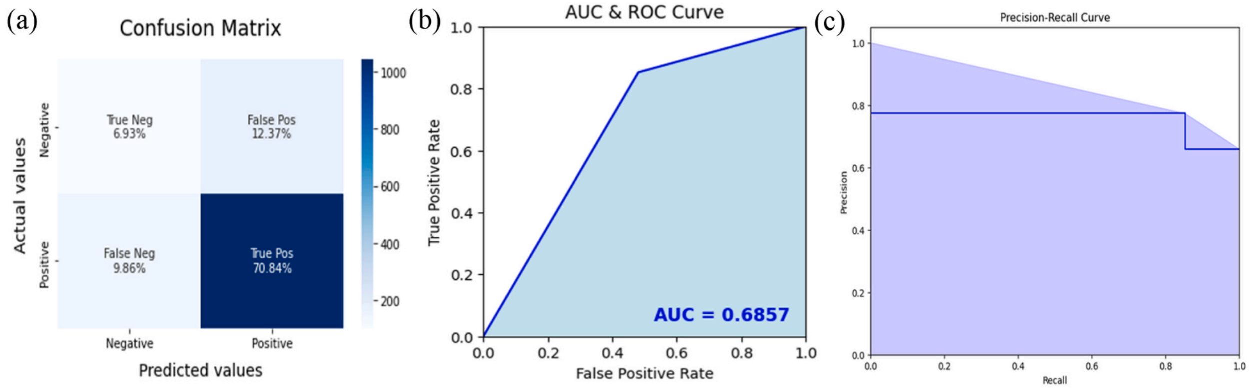

Fig. 9. Performance Metrics for K-Nearest Neighbor (KNN) classifier. (a) Confusion matrix, (b) AUC & ROC curve and (c) Precision-Recall curve.

Table 4
Experiment Results for K-Nearest Neighbor (KNN) classifier.

	Precision	Recall	F1-score	Accuracy	AUC	AUPR
Negative	41 %	36 %	38 %	78 %	0.6857	0.8616
Positive	85 %	88 %	86 %			

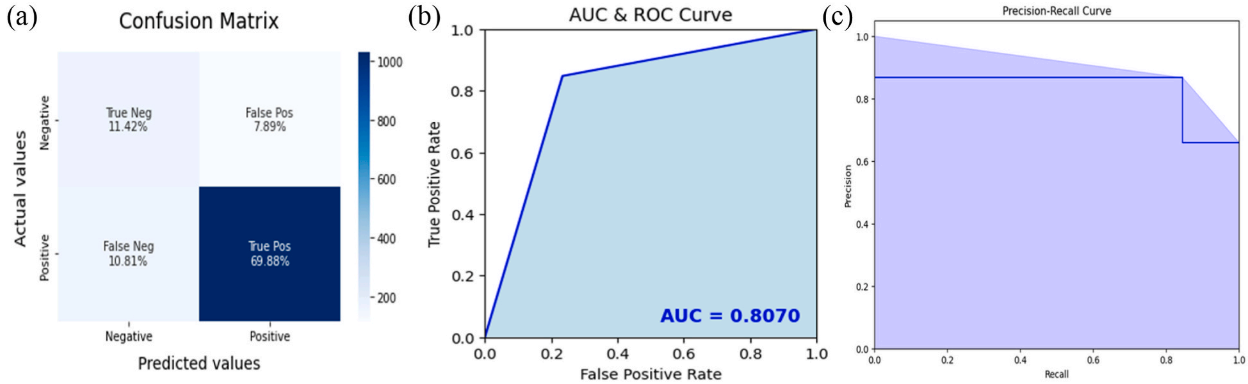

Fig. 10. Performance Metrics for Decision Trees (DTs) classifier. (a) Confusion matrix, (b) AUC & ROC curve and (c) Precision-Recall curve.

Table 5
Experiment Results for Decision Trees (DTs) classifier.

	Precision	Recall	F1-score	Accuracy	AUC	AUPR
Negative	51 %	59 %	55 %	81 %	0.8070	0.9073
Positive	90 %	87 %	88 %			

instant rises or falls are encountered. For investors, analysis of individual behavior variability on different company shares gives insight into the market and helps them to make better investments with the help of Twitter sentiment analysis. By utilizing sentiment analysis on tweets about finance, the research obtained valuable insights into the sentiment and perceptions of the general public regarding a wide range of stock market companies. By employing this methodology, one can effectively assess the prevailing sentiment of the investing public. The quantitative assessment of public attention and interest can be obtained by analyzing the volume of tweets about various company shares. The objective of the methodology was to ascertain and analyze the fluctuations in investor mood and levels of interest over a certain period.

In our case, it was determined that there was dominant positivity in tweets about company shares between the dates of 21.11.2022 and 25.11.2022. Only two of the shares, HEKTS and SASA, had dominant negativity in tweets as their prices were decreasing during the above-mentioned dates. Yet, those two shares had less decrease in prices during that period which caused misclassification of the models, especially for negative classes. The most important reason why these stock prices decreased less in the relevant periods is the

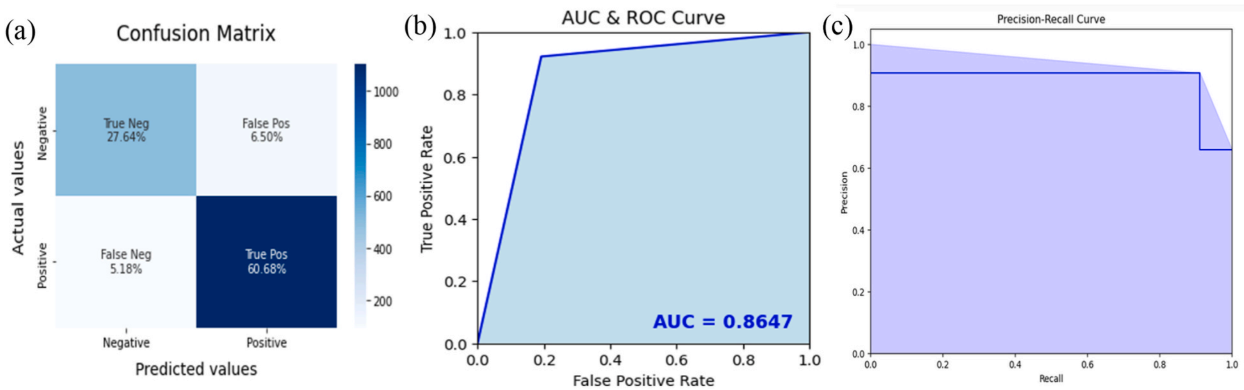

Fig. 11. Performance Metrics for Multilayer Perceptual (MLP) classifier. (a) Confusion matrix, (b) AUC & ROC curve and (c) Precision-Recall curve.

Table 6
Experiment Results for Multilayer Perceptual (MLP) classifier.

	Precision	Recall	F1-score	Accuracy	AUC	AUPR
Negative	84 %	81 %	83 %	88 %	0.8647	0.9383
Positive	90 %	92 %	91 %			

Table 7
Summary of experimental results.

	Classes	Precision	Recall	F1-score	Accuracy	AUC	AUPR
NB	Negative	52 %	88 %	65 %	68 %	0.7297	0.8794
	Positive	90 %	68 %	71 %			
LR	Negative	83 %	81 %	82 %	88 %	0.8506	0.9316
	Positive	90 %	91 %	91 %			
SVM	Negative	85 %	82 %	84 %	89 %	0.8729	0.9415
	Positive	91 %	93 %	92 %			
KNN	Negative	41 %	36 %	38 %	78 %	0.6857	0.8616
	Positive	85 %	88 %	86 %			
DTs	Negative	51 %	59 %	55 %	81 %	0.8070	0.9073
	Positive	90 %	87 %	88 %			
MLP	Negative	84 %	81 %	83 %	88 %	0.8647	0.9383
	Positive	90 %	92 %	91 %			

reluctance of investors to sell their shares based on past price movements. Both stocks have yielded an average annual return of over 300 % over the last 5 years. Considering the ongoing large investments of these companies, it is natural that investors do not want to sell their stocks. Thus, this study suggests that for a comprehensive analysis, a new analysis should be made by including data from different periods and data from other social media sites or platforms, information obtained from news sources, and various economic indicators, as well. During the observation period, it was found that the majority of users on Twitter.com were Turkish-speaking and based in Türkiye. Future analyses must incorporate considerations of both location and language to expand the investors' profile. It would be also possible to increase sentiment analysis accuracy with parameter optimization on a larger and more balanced dataset for training purposes of the ML models. Working with cleaner data by changing the pre-processing steps for the data gathered from the online environment can also help to get better results.

6. Limitations

Disclosure of the study's constraints, including the intricacy of market dynamics and external influences, enhances its transparency. The act of identifying gaps in knowledge promotes an ongoing investigation into the intricate correlation that exists between sentiment expressed on social media platforms and the fluctuations of the stock market. This study utilizes tweets about financial matters as the major source of data. Although this analysis offers a glimpse into the prevailing public opinion, it is important to acknowledge that it may not encompass the entirety of pertinent information or sentiments conveyed by alternative mediums.

The present study acknowledges the aforementioned restriction and places emphasis on the potential variability in the accuracy of sentiment labels. Further investigation could be conducted to examine more sophisticated sentiment analysis approaches to enhance precision. Additionally, our study recognizes the impact of external factors, such as economic occurrences, policy modifications, or worldwide market patterns, on stock valuations. Nevertheless, it is possible that this approach does not fully encompass the intricacies

Table 8
Opening and closing prices of BIST30 Firms and tweets posted between November 21, 2022, and November 25, 2022.

	Firms	(21.11.2022) Opening Price	(25.11.2022) Closing Price	21–25.11.2022 # of Tweets
1	Akbank TAS (AKBNK)	16,41	16,82	1200
2	Aksa Enerji Uretim AS (AKSEN)	48,66	53,05	1046
3	Arcelik AS (ARCLK)	85,20	91,90	702
4	Aselsan Elektronik Sanayi ve Ticaret AS (ASELS)	40,04	48,20	2588
5	BIM Birlesik Magazalar AS (BIMAS)	131,80	132,30	515
6	Emlak Konut Gayrimenkul Yatirim Ortakligi AS (EKGYO)	5,53	5,78	898
7	Eregli Demir ve Celik Fabrikalari TAS (EREGL)	37,06	41,88	4223
8	Ford Otomotiv Sanayi AS (FROTO)	436,40	447,60	1020
9	Gubre Fabrikalari TAS (GUBRF)	162,00	187,40	452
10	Haci Omer Sabanci Holding AS (SAHOL)	36,60	40,20	1343
11	Hektas Ticaret AS (HEKTS)	37,40	34,98	7377
12	Kardemir Karabuk Demir Celik Sanayi ve Ticaret AS (KRDMD)	12,89	14,82	1639
13	Koc Holding AS (KCHOL)	63,25	70,40	1061
14	Koza Altin Isletmeleri AS (KOZAL)	256,50	341,70	1448
15	Koza Anadolu Metal Madencilik Isletmeleri AS (KOZAA)	37,70	44,90	908
16	Pegasus Hava Tasimaciligi AS (PGSUS)	346,30	380,80	977
17	Petkim Petrokimya Holding AS (PETKM)	15,14	15,29	2146
18	Sasa Polyester Sanayi AS (SASA)	135,90	130,20	10000
19	TAV Havalimanlari Holding (TAVHL)	78,44	86,35	416
20	Tekfen Holding AS (TKFEN)	37,46	40,66	537
21	Tofas Turk Otomobil Fabrikasi AS (TOASO)	126,00	134,10	796
22	Turkcell Iletisim Hizmetleri AS (TCELL)	31,56	34,52	728
23	Turkiye Petrol Rafinerileri AS (TUPRS)	426,30	473,70	1723
24	Turk Hava Yolları AO (THYAO)	106,70	113,80	2190
25	Turk Telekomunikasyon AS (TTKOM)	14,64	17,10	1060
26	Turkiye Garanti Bankasi AS (GARAN)	25,50	26,38	1477
27	Turkiye Is Bankası AS (ISCTR)	9,64	10,25	777
28	Turkiye Sise ve Cam Fabrikalari AS (SISE)	34,84	39,96	1678
29	Vestel Elektronik Sanayi ve Ticaret AS (VESTL)	43,50	50,60	856
30	Yapi ve Kredi Bankasi AS (YKBNK)	10,01	11,45	1123

Table 9
Weekly sentiment predictions of top-mentioned company tweets using machine learning approaches.

		# of Total Tweets	NB		LR		SVM		KNN		DTs		MLP	
			Pos.	Neg.	Pos.	Neg.	Pos.	Neg.	Pos.	Neg.	Pos.	Neg.	Pos.	Neg.
1	ASELS	2588	909	1679	1745	843	1789	799	2033	555	1674	914	1844	744
2	EREGL	4223	1343	2880	2756	1467	2944	1279	3179	1044	2683	1540	2843	1380
3	HEKTS	7377	1906	5471	4381	2996	3167	4210	5868	1509	4903	2474	4987	2390
4	SASA	10000	2473	7527	6141	3859	5872	4128	7730	2270	7138	2862	6104	3896

inherent in these relationships. Further investigation is warranted to explore the intricate relationship between sentiment expressed on social media platforms and external market dynamics.

Moreover, our study primarily examines the relationship between tweet volume and sentiment, but it does not extensively explore company-specific elements that could potentially influence stock changes. Various factors, such as the release of earnings reports, changes in management, or the introduction of new products, have the potential to exert a substantial influence on stock prices, hence necessitating a more in-depth examination.

7. Conclusion

This study aims to contribute to the literature by examining the tweets in the Turkish language about BIST30 company shares in Borsa Istanbul with sentiment analysis to investigate public mood for stock market prediction. Tweets consisting of hashtags "#Borsaistanbul, #Bist, #Bist30, #Bist100″ were imported from November 7, 2022, to November 15, 2022, via the MAXQDA 2020 program. Orange Data Mining program was used to label the polarities of tweets as positive, negative, and neutral with the multilingual sentiment method provided in the program. Then, 9076 samples out of 17189 tweets were selected as negative and positive for machine learning training and testing purposes by implementing applications in Python 3.6 with the sklearn library. Train-test split using the 80-20 rule was applied to the dataset. 80 % of the data was used for training and the remaining 20 % was used for testing (10 %) and validation (10 %). We compared machine learning approaches with tf-idf vectorizers. As shown in Table 7, the highest accuracy rate was achieved by the Support Vector Machines and Multilayer Perceptron classifier AUC values of 0.8729 and 0.8647 and AUPR values of 0.9415 and 0.9383 respectively. Other algorithms obtained approximately a 78,75 % accuracy rate.

The literature review has revealed that there are many studies making use of machine learning approaches in sentiment analysis in

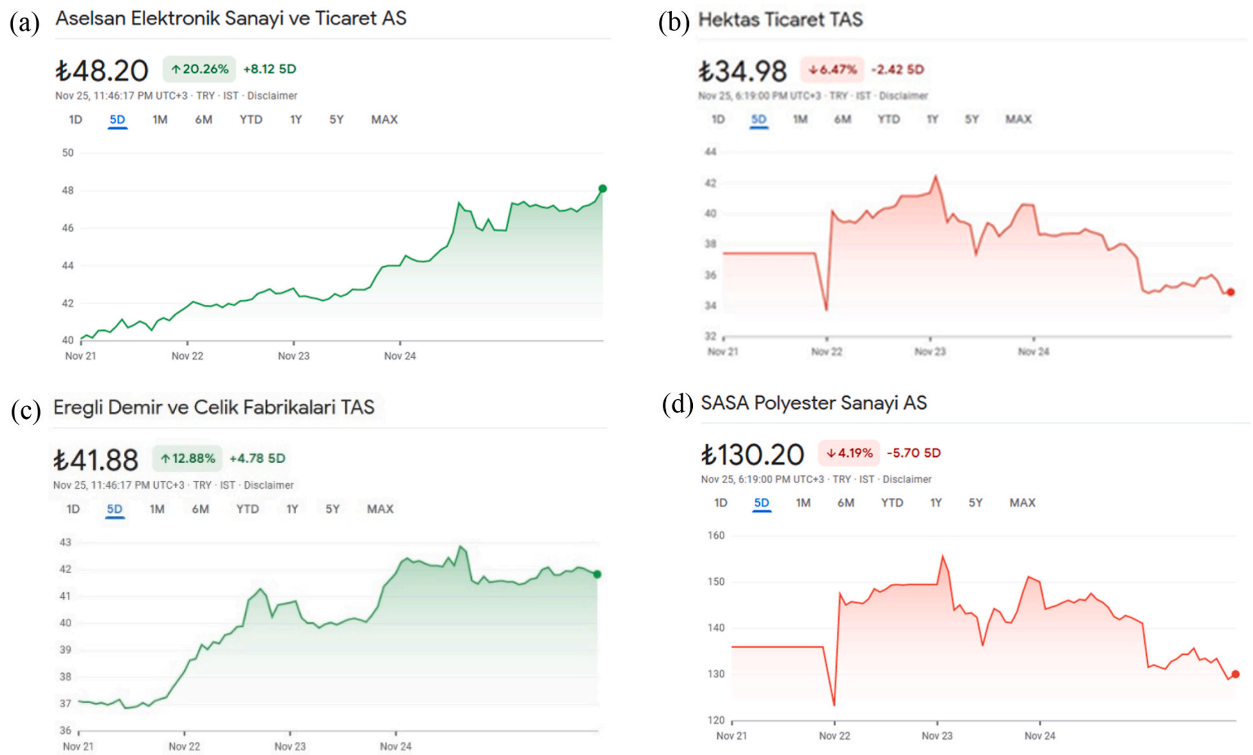

Fig. 12. Graphs of price index directions of four firms between November 21, 2022, and November 25, 2022. (a) Aselsan Elektronik Sanayi ve Ticaret AS price index direction. (b) Hektas Ticaret AS price index direction. (c) Eregli Demir ve Celik Fabrikalari TAS price index direction. (d) SASA Polyester Sanayi AS price index direction.

the Turkish language focusing on different fields such as sentiment analysis of Turkish political news [23], classification of Turkish movie reviews [24], Turkish Twitter posts sentiment analysis [49] and more [25,26,47,48,50–52,55]. However, even though there have been studies on stock market predictions with sentiment analysis in other languages [5], one study [72] has been conducted in Turkish within the scope of stock market prediction to the best of our knowledge. Our main motivation behind this study is to apply sentiment analysis on financial-related tweets in Turkish by integrating hybrid lexicon-based with machine learning methods. However, the transmission of information within an online social networking environment occurs rapidly and effortlessly, resulting in a leveling impact on the investing methods and emotional states of traders. Thus, for future academic inquiry, a closer examination of the connection between information and public emotional states may be warranted.

Further investigation could be conducted to examine more sophisticated sentiment analysis models and natural language processing techniques to improve the precision and detail of sentiment classification. One potential approach is the utilization of context-aware sentiment analysis techniques to more effectively capture the intricate nuances present in tweets about financial matters. Furthermore, the incorporation of multimodal data, such as images or links provided inside tweets, has the potential to enhance the depth and breadth of information available for sentiment analysis. The investigation of the relationship between multimedia material and its influence on public attitude, as well as the subsequent effects on stock market behavior, presents a promising area for future academic inquiry. A full comprehension of the influence of various online communities on stock market perceptions can be achieved by broadening the investigation to encompass several social media platforms and by comparing sentiment patterns across these platforms.

Addressing these limitations and exploring these avenues for further research will contribute to the ongoing discourse on the interplay between social media sentiment and stock market dynamics. In our future work, sentiment analysis with machine learning approaches used in this work can be expanded to develop better sentiment analysis using deep learning approaches.

Funding statement

This research received no specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

Data will not be made available in accordance with the Twitter Privacy Policy.

Additional information

No additional information is available for this paper.

CRediT authorship contribution statement

Handan Cam: Writing – review & editing, Writing – original draft, Supervision, Methodology, Formal analysis, Data curation, Conceptualization. **Alper Veli Cam:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Formal analysis, Data curation, Conceptualization. **Ugur Demirel:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Sana Ahmed:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A. Go, L. Huang, R. Bhayani, Twitter sentiment analysis, Entropy 17 (2009) 252.
[2] O. Appel, F. Chiclana, J. Carter, H. Fujita, A hybrid approach to sentiment analysis, in: 2016 IEEE Congress on Evolutionary Computation (CEC), IEEE, 2016, pp. 4950–4957.
[3] A. Nandi, P. Sharma, Comparative study of sentiment analysis techniques, in: Interdisciplinary Research in Technology and Management, CRC Press, 2021, pp. 456–460.
[4] B. Pang, L. Lee, Opinion mining and sentiment analysis, Foundations and Trends® in Information Retrieval 2 (2008) 1–135.
[5] J. Bollen, H. Mao, X. Zeng, Twitter mood predicts the stock market, Journal of Computational Science 2 (2011) 1–8.
[6] N. Pervan, H.Y. Keleş, Sentiment Analysis Using a Random Forest Classifier on Turkish Web Comments, vol. 59, Communications Faculty of Sciences University of Ankara Series A2-A3 Physical Sciences and Engineering, 2017, pp. 69–79.
[7] About (n.d.), https://www.maxqda.com/about. (Accessed 12 September 2023).
[8] Orange (n.d. https://orangedatamining.com/. (Accessed 12 September 2023).
[9] S. Cheng, Y. Jin, S.P. Harrison, C. Quilodrán-Casas, I.C. Prentice, Y.-K. Guo, R. Arcucci, Parameter flexible wildfire prediction using machine learning techniques: forward and inverse modelling, Rem. Sens. 14 (2022) 3228.
[10] S. Bayram, H. Çıtakoğlu, Modeling monthly reference evapotranspiration process in Turkey: application of machine learning methods, Environ. Monit. Assess. 195 (2023) 67.
[11] V. Demir, H. Citakoglu, Forecasting of solar radiation using different machine learning approaches, Neural Comput. Appl. 35 (2023) 887–906.
[12] H. Citakoglu, Ö. Coşkun, Comparison of hybrid machine learning methods for the prediction of short-term meteorological droughts of Sakarya Meteorological Station in Turkey, Environ. Sci. Pollut. Control Ser. 29 (2022) 75487–75511.
[13] D. Maynard, A. Funk, Automatic detection of political opinions in tweets, in: Extended Semantic Web Conference, Springer, 2011, pp. 88–99.
[14] W. Medhat, A. Hassan, H. Korashy, Sentiment analysis algorithms and applications: a survey, Ain Shams Eng. J. 5 (2014) 1093–1113.
[15] B. Pang, L. Lee, S. Vaithyanathan, Thumbs up? Sentiment Classification Using Machine Learning Techniques, 2002 arXiv Preprint Cs/0205070.
[16] X. Bai, Predicting consumer sentiments from online text, Decis. Support Syst. 50 (2011) 732–742.
[17] J.V. Lochter, R.F. Zanetti, D. Reller, T.A. Almeida, Short text opinion detection using ensemble of classifiers and semantic indexing, Expert Syst. Appl. 62 (2016) 243–249.
[18] A. Hassan, A. Mahmood, Deep learning approach for sentiment analysis of short texts, in: 2017 3rd International Conference on Control, Automation and Robotics (ICCAR), IEEE, 2017, pp. 705–710.
[19] M.R. Saleh, M.T. Martín-Valdivia, A. Montejo-Ráez, L. Ureña-López, Experiments with SVM to classify opinions in different domains, Expert Syst. Appl. 38 (2011) 14799–14804.
[20] R.M. Duwairi, Sentiment analysis for dialectical Arabic, in: 2015 6th International Conference on Information and Communication Systems (ICICS), IEEE, 2015, pp. 166–170.
[21] Z. Zhang, Q. Ye, Z. Zhang, Y. Li, Sentiment classification of Internet restaurant reviews written in Cantonese, Expert Syst. Appl. 38 (2011) 7674–7682.
[22] H. Zhang, W. Gan, B. Jiang, Machine learning and lexicon based methods for sentiment classification: a survey, in: 2014 11th Web Information System and Application Conference, IEEE, 2014, pp. 262–265.
[23] M. Kaya, G. Fidan, I.H. Toroslu, Sentiment analysis of Turkish political news, in: 2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, IEEE, 2012, pp. 174–180.
[24] F. Akba, A. Uçan, E.A. Sezer, H. Sever, Assessment of feature selection metrics for sentiment analyses: Turkish movie reviews, in: 8th European Conference on Data Mining, 2014, pp. 180–184.
[25] B. Ciftci, M.S. Apaydin, A deep learning approach to sentiment analysis in Turkish, in: 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), IEEE, 2018, pp. 1–5.
[26] T. Parlar, S.A. Özel, F. Song, QER: a new feature selection method for sentiment analysis, Human-Centric Computing and Information Sciences 8 (2018) 1–19.
[27] Y. Santur, Sentiment analysis based on gated recurrent unit, in: 2019 International Artificial Intelligence and Data Processing Symposium (IDAP), IEEE, 2019, pp. 1–5.
[28] U.U. Acikalin, B. Bardak, M. Kutlu, Turkish sentiment analysis using bert, in: 2020 28th Signal Processing and Communications Applications Conference (SIU), IEEE, 2020, pp. 1–4.
[29] I. Kurniawati, H.F. Pardede, Hybrid method of information gain and particle swarm optimization for selection of features of SVM-based sentiment analysis, in: 2018 International Conference on Information Technology Systems and Innovation (ICITSI), IEEE, 2018, pp. 1–5.
[30] A. Rane, A. Kumar, Sentiment classification system of Twitter data for US airline service analysis, in: 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), IEEE, 2018, pp. 769–773.
[31] V.L.S. Lee, K.H. Gan, T.P. Tan, R. Abdullah, Semi-supervised learning for sentiment classification using small number of labeled data, Procedia Comput. Sci. 161 (2019) 577–584.
[32] S.A. El Rahman, F.A. AlOtaibi, W.A. AlShehri, Sentiment analysis of twitter data, in: 2019 International Conference on Computer and Information Sciences (ICCIS), IEEE, 2019, pp. 1–4.
[33] P. Ray, A. Chakrabarti, A mixed approach of deep learning method and rule-based method to improve aspect level sentiment analysis, Appl. Comput. Inform. 18 (1/2) (2022) 163–178.
[34] A. Hasell, Shared emotion: the social amplification of partisan news on Twitter, Digital Journalism 9 (2021) 1085–1102.

[35] S. Lazarus, M. Button, A. Adogame, Advantageous comparison: using Twitter responses to understand similarities between cybercriminals (“Yahoo Boys”) and politicians (“Yahoo men”), Heliyon 8 (2022).
[36] P. Vyas, M. Reisslein, B.P. Rimal, G. Vyas, G.P. Basyal, P. Muzumdar, Automated classification of societal sentiments on Twitter with machine learning, IEEE Transactions on Technology and Society 3 (2021) 100–110.
[37] M.M. Alam, H. Malik, M.I. Khan, T. Pardy, A. Kuusik, Y. Le Moullec, A survey on the roles of communication technologies in IoT-based personalized healthcare applications, IEEE Access 6 (2018) 36611–36631.
[38] N. Chintalapudi, G. Battineni, F. Amenta, Sentimental analysis of COVID-19 tweets using deep learning models, Infect. Dis. Rep. 13 (2021) 329–339.
[39] H. Kaur, S.U. Ahsaan, B. Alankar, V. Chang, A proposed sentiment analysis deep learning algorithm for analyzing COVID-19 tweets, Inf. Syst. Front (2021) 1–13.
[40] F. Rustam, M. Khalid, W. Aslam, V. Rupapara, A. Mehmood, G.S. Choi, A performance comparison of supervised machine learning models for Covid-19 tweets sentiment analysis, PLoS One 16 (2021), e0245909.
[41] M. Qorib, T. Oladunni, M. Denis, E. Ososanya, P. Cotae, Covid-19 vaccine hesitancy: text mining, sentiment analysis and machine learning on COVID-19 vaccination Twitter dataset, Expert Syst. Appl. 212 (2023), 118715.
[42] M.M. Rahman, M.N. Islam, Exploring the performance of ensemble machine learning classifiers for sentiment analysis of COVID-19 tweets, in: Sentimental Analysis and Deep Learning: Proceedings of ICSADL 2021, Springer, 2021, pp. 383–396.
[43] M. Cliche, BB_twtr at SemEval-2017 Task 4: Twitter Sentiment Analysis with CNNs and LSTMs, 2017 arXiv Preprint arXiv:1704.06125.
[44] S. Kamis, D. Goularas, Evaluation of deep learning techniques in sentiment analysis from twitter data, in: 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML), IEEE, 2019, pp. 12–17.
[45] R.D. Desai, Sentiment analysis of Twitter data, in: 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), IEEE, 2018, pp. 114–117.
[46] S. Al-Hadhrami, N. Al-Fassam, H. Benhidour, Sentiment analysis of English tweets: a comparative study of supervised and unsupervised approaches, in: 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS), IEEE, 2019, pp. 1–5.
[47] H. Nizam, S.S. Akın, Sosyal medyada makine öğrenmesi ile duygu analizinde dengeli ve dengesiz veri setlerinin performanslarının karşılaştırılması, XIX. Türkiye'de İnternet Konferansı. 1 (2014).
[48] M. Meral, B. Diri, Sentiment analysis on twitter, in: 2014 22nd Signal Processing and Communications Applications Conference (SIU), IEEE, 2014, pp. 690–693.
[49] Ö. Çoban, B. Özyer, G.T. Özyer, Sentiment analysis for Turkish Twitter feeds, in: 2015 23nd Signal Processing and Communications Applications Conference (SIU), IEEE, 2015, pp. 2388–2391.
[50] O. Kaynar, Y. Görmez, M. Yıldız, A. Albayrak, Makine öğrenmesi yöntemleri ile duygu analizi, in: International Artificial Intelligence and Data Processing Symposium (IDAP'16), 2016, p. 241.
[51] C. Catal, M. Nangir, A sentiment classification model based on multiple classifiers, Appl. Soft Comput. 50 (2017) 135–141.
[52] R. Velioğlu, T. Yıldız, S. Yıldırım, Sentiment analysis using learning approaches over emojis for Turkish tweets, in: 2018 3rd International Conference on Computer Science and Engineering (UBMK), IEEE, 2018, pp. 303–307.
[53] G.M. Demirci, Ş.R. Keskin, G. Doğan, Sentiment analysis in Turkish with deep learning, in: 2019 IEEE International Conference on Big Data (Big Data), IEEE, 2019, pp. 2215–2221.
[54] H.A. Shehu, S. Tokat, A hybrid approach for the sentiment analysis of Turkish Twitter data, in: The International Conference on Artificial Intelligence and Applied Mathematics in Engineering, Springer, 2020, pp. 182–190.
[55] T. Parlar, E. Saraç, S.A. Özel, Comparison of feature selection methods for sentiment analysis on Turkish Twitter data, in: 2017 25th Signal Processing and Communications Applications Conference (SIU), IEEE, 2017, pp. 1–4.
[56] A. Esuli, F. Sebastiani, Sentiwordnet: a publicly available lexical resource for opinion mining, in: Proceedings of the Fifth International Conference on Language Resources and Evaluation, LREC'06), 2006.
[57] A. Baloglu, M.S. Aktas, An automated framework for mining reviews from blogosphere, International Journal of Advances in Internet Technology 3 (2010) 234–244.
[58] X. Ding, B. Liu, P.S. Yu, A holistic lexicon-based approach to opinion mining, in: Proceedings of the 2008 International Conference on Web Search and Data Mining, 2008, pp. 231–240.
[59] A. John, A. John, R. Sheik, Context deployed sentiment analysis using hybrid lexicon, in: 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), IEEE, 2019, pp. 1–5.
[60] O. Bilgin, Ö. Çetinoğlu, K. Oflazer, Building a wordnet for Turkish, Rom. J. Inf. Sci. Technol. 7 (2004) 163–172.
[61] A.G. Vural, B.B. Cambazoglu, P. Senkul, Z.O. Tokgoz, A framework for sentiment analysis in Turkish: application to polarity detection of movie reviews in Turkish, in: Computer and Information Sciences III, Springer, 2013, pp. 437–445.
[62] E.S. Akgül, C. Ertano, D. Banu, Sentiment analysis with twitter, Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi. 22 (2016) 106–110.
[63] R. Dehkharghani, Y. Saygin, B. Yanikoglu, K. Oflazer, SentiTurkNet: a Turkish polarity lexicon for sentiment analysis, Comput. Humanit. 50 (2016) 667–685.
[64] R. Dehkharghani, B. Yanikoglu, Y. Saygin, K. Oflazer, Sentiment analysis in Turkish at different granularity levels, Nat. Lang. Eng. 23 (2017) 535–559.
[65] N. Öztürk, S. Ayvaz, Sentiment analysis on Twitter: a text mining approach to the Syrian refugee crisis, Telematics Inf. 35 (2018) 136–147.
[66] G. Yurtalan, M. Koyuncu, Ç. Turhan, A polarity calculation approach for lexicon-based Turkish sentiment analysis, Turk. J. Electr. Eng. Comput. Sci. 27 (2019) 1325–1339.
[67] Sentiment Analysis, (n.d.). https://orangedatamining.com/widget-catalog/text-mining/sentimentanalysis/(accessed November 23, 2022).
[68] B. Ohana, B. Tierney, Sentiment classification of reviews using SentiWordNet, Proceedings of IT&T. 8 (2009).
[69] M. Govindarajan, Sentiment analysis of movie reviews using hybrid method of naive bayes and genetic algorithm, Int. J. Adv. Comput. Res. 3 (2013) 139.
[70] M. Rumelli, D. Akkuş, Ö. Kart, Z. Isik, Sentiment analysis in Turkish text with machine learning algorithms, in: 2019 Innovations in Intelligent Systems and Applications Conference (ASYU), IEEE, 2019, pp. 1–5.
[71] B. ErSahin, O. Aktaş, D. Kilinc, M. Erşahin, A hybrid sentiment analysis method for Turkish, Turk. J. Electr. Eng. Comput. Sci. 27 (2019) 1780–1793.
[72] M.U. Şimşek, S. Özdemir, Analysis of the relation between Turkish twitter messages and stock market index, in: 2012 6th International Conference on Application of Information and Communication Technologies (AICT), IEEE, 2012, pp. 1–4.
[73] Twitter Data, (n.d.). https://www.maxqda.com/help-mx20/import/importing-twitter-data (accessed November 23, 2022).
[74] Preprocess Text, (n.d.). https://orangedatamining.com/widget-catalog/text-mining/preprocesstext/(accessed November 23, 2022).
[75] U.Ö. Osmanoğlu, O.N. Atak, K. Çağlar, H. Kayhan, T.C. Can, Sentiment analysis for distance education course materials: a machine learning approach, Journal of Educational Technology and Online Learning 3 (2020) 31–48.
[76] Ö. Çelik, A.F. Aslan, Gender prediction from social media comments with artificial intelligence, Sakarya University Journal of Science 23 (2019) 1256–1264.
[77] Precision-Recall, (n.d.). https://scikit-learn.org/stable/auto_examples/model_selection/plot_precision_recall.html (accessed September 12, 2023).
[78] BIST 30 (XU030), (n.d.). https://tr.investing.com/indices/ise-30-historical-data (accessed November 23, 2022).