

TURKISH SENTIMENT ANALYSIS FOR OPEN AND DISTANCE EDUCATION SYSTEMS

Zeliha ERGUL AYDIN

ORCID: 0000-0002-7108-8930
Department of Industrial Engineering
Eskisehir Technical University
Eskisehir, TURKEY

Dr. Zehra KAMISLI OZTURK

ORCID: 0000-0003-3156-6464
Department of Industrial Engineering
Eskisehir Technical University
Eskisehir, TURKEY

Dr. Zeynep Idil ERZURUM CICEK

ORCID: 0000-0001-9641-8935
Department of Industrial Engineering
Eskisehir Technical University
Eskisehir, TURKEY

Received: 15/06/2020 **Accepted:** 11/12/2020

ABSTRACT

Students' opinions are the most essential source to enhance the quality of education and educational services in Open and Distance education (ODE) Systems. How to access and analyze students' real opinions is a problem for ODE institutions. The purpose of the present study is to conduct a sentiment analysis (SA) on the collected Turkish tweets about an ODE system to monitor students' opinions and sentiments about the system. Firstly, the related 63699 tweets about the ODE system are gathered and analyzed. Later, pre-processing is applied to the dataset. Sentence-based SA is performed with the data provided. The dataset is vectorized using two vector space models to test four classifiers which are Support Vector Machines, K-Nearest Neighbor, Logistic Regression (LR), and Artificial Neural Networks. F-score values obtained with these classifiers are evaluated, and the results are discussed. LR classifier gives the best F-score values with %75 for each vector space model. Through the SA results, students' dissatisfaction, appreciation, and concerns will be learned quickly by the university administration to develop strategies that will increase the quality of education and educational services.

Keywords: Sentiment analysis, machine learning, open and distance education system, natural language processing, social media, Twitter.

INTRODUCTION

Sentiment analysis (SA) - How people feel about a topic (e.g., their sentiment) can be just as critical as identifying the topic itself (Barbier & Liu, 2011). These feelings can be extracted from a written or spoken language using Natural Language Processing (NLP) techniques and machine learning. SA has raised growing interest, both within the scientific community, leading to many exciting challenges, as well as in the business world, due to the remarkable benefits to be had from financial and political forecasting, user profiling and community detection, computational advertising, and dialogue systems. However, mining opinions and sentiments from multimodal resources (texts, images, videos, audio-recordings, etc.) is a challenging task because it requires a deep understanding of the explicit and implicit, regular and irregular, features (linguistic, visual, or audio) of a resource (Dragoni et al., 2018).

In the literature, SA is alternatively named as opinion mining, sentiment mining, opinion extraction, subjectivity analysis, and artificial emotional intelligence (emotion AI). It will be useful to point out a common complication. In the literature, SA and opinion mining are often used interchangeably. For instance, as given by Sangam and Shinde (2019), “the task of analyzing the opinion, sentiment, and subjectivity computationally are known as opinion mining,” and it is also called as SA. However, According to Pozzi et al. (2016), an opinion is more of a person’s detailed view about something, whereas a sentiment is more of a feeling.

SA applications’ primary purpose is to analyze people’s sentiments or opinions toward entities such as topics, events, individuals, issues, services, products, organizations, and their attributes (Liu, 2012). In the Social Web context, through the popularization of the platforms, which provide access to large amounts of subjective data, trying to automate the understanding of people’s opinions about theme, service, or product has been a key factor in the decision-making process (de Oliveira et al., 2021). In recent years, SA is used to analyze consumer behavior, predict revenue, and classify consumer reviews in several fields like healthcare, political events, finance, marketing, and education.

SA was used in healthcare to investigate the relation between neighborhood characteristics and obesity, diabetes (Nguyen et al., 2017) and Alzheimer’s Disease Stigma on Twitter (Oscar et al., 2017). Especially, during election periods, SA is a powerful technique for candidates’ promotional campaign decisions (Opuszko et al., 2018), (Ahmed et al., 2016) and prediction of selection results (Bansal & Srivastava, 2018), (Kušen & Strembeck, 2018), (Jaidka et al., 2019). On the marketing side, various shopping sites like Amazon and Flipkart take feedback from the customers, which will help them make proper decisions to improve the quality of their services (Sangam & Shinde, 2019). Stock price forecasting (Nguyen et al., 2015), detection of the determinants of Bitcoin Prices (Georgoula et al., 2015) and measuring inflation expectations (Goloshchapova & Andreev, 2017) are some of the examples of finance SA applications.

SA is also a powerful technique used in educational institutions. SA applications in education can be classified as instruction evaluation, institutional decision/policy making, intelligent information/learning systems enhancement and assignment evaluation, and feedback improvement according to the different task types that sentiment analysis tools have served in the domain of education (Dolianiti et al., 2019). It is not enough for educational institutions to consider students’ education, registration, and course information to improve education quality in today’s world. Open and distance education (ODE) represents approaches that focus on opening access to education and training provision, freeing learners from time and place constraints, and offering flexible learning opportunities to individuals and groups of learners. Open and distance learning is one of the most rapidly growing fields of education. Its potential impact on all education delivery systems has been dramatically accentuated by developing internet-based information technologies, particularly the World Wide Web (Kamisli Ozturk, 2012). In the ODE systems, considering the high number of students, all kinds of data from the students should be evaluated to improve the quality of education and educational services continuously. Students’ opinions about educational institutions and their services play an essential role in enhancing quality. One of the biggest challenges faced by ODE institutions is accessing and analyzing students’ real opinions and sentiments about the ODE system. Students’ social media data is one of the best data sources to understand students’ opinions about the institution because social media allows students to express their ideas freely. Twitter, Facebook, Google +, and Instagram can be given as examples of social media platforms where these opinions can be collected. Therefore, SA can be used to understand the students’ opinions from these platforms.

This study, which is based on students' feedback improvement, aims to present a SA model to extract the opinions of an ODE institution's students in Turkey. The SA model classifies students' tweets about their institution into sentiments (positive, negative, or neutral) with different classifiers and vector space models. This study aims to use educational mining data to help managers in the decision-making process to achieve this goal. As it's given by Sorour et al. (2020), some levels in the organizations require real-time information for detection and correction of nonconformance to goal activities (Sorour et al., 2020). In this study, sentiment analysis studies have been conducted to understand students' opinions regarding the distance education system. Such an analysis provides the opportunity to analyze the text data to summarize the students' views about the educational institution they are affiliated with and identify positive and negative comments about the system. With this approach, it helps the relevant decision-making process for the whole system in distance education. From this point of view, the present study aims to answer the following research questions:

RQ1: How can we automatically classify the sentiment of the students' Turkish tweets about the ODE system for understanding the students' real sentiments about the system?

RQ2: Can different vector space models increase the success of sentiment classification?

The rest of this paper is structured as following: SA application in students' feedback improvement has been widely evaluated in Literature Review Section, the structure and principles of SA have been presented in the Sentiment Analysis Section, the methodology has been given in Methodology Section, the results and limitations of the presented study have been discussed in Discussions Section, and the conclusion and future works have been given Conclusions Section.

LITERATURE REVIEW

As Ali and Abdel-Haq (2020) mentioned in their study, using artificial intelligence in education has been the subject of enormous debates for the past 30. Generally speaking, we can say that artificial intelligence applications offer both potential pedagogical opportunities and opportunities for improving distance education systems to support students throughout their lifecycle. SA, one of the powerful tools of artificial intelligence, is crucial in education where student's feedback is required to be assessed for the learning (Kim & Calvo, 2010). In the literature, there are many studies about student feedback. These studies have collected data from opinion surveys, educational portals/websites/blogs, and social media.

Opinion Surveys

To better understand the influences and opportunities to increase student satisfaction, Thomas and Galambos (2004) used regression and decision tree analysis on data, collected from a student opinion survey at a public university. Aung and Myo (2017) gathered students' comments from the Faculty Evaluation Survey for the University of Computer Studies and applied a lexicon-based SA approach. Such feedbacks are generally gathered at the end of the semester with the use of survey forms. However, as Ullah (2016) mentioned, this process is very tedious, slow, and time-consuming.

Educational Portals, Web Sites and Blogs

Wen et al. (2014) explored collective mining sentiment from forum posts of Coursera, Massive Open Online Course platform, to monitor students' trending opinions towards the course and significant course tools, such as lecture and peer-assessment. Furthermore, Tucker et al. applied SA on student-generated textual data (e.g., online discussion forums) existing in MOOCs to quantify their impact on student performance and learning outcomes. Kechaou et al. (2011) applied SA to examine the nature and the structure of web forums and e-learning blogs to provide a better understanding of users' opinions regarding the e-learning system for the sake of its improvement. In their paper, Song et al. (2017) applied SA to help e-learning systems to know the users' opinions on the course-ware, the teachers, the charge or something else of the e-learning system, and to help the developers improve the services. Since system administrators control education portals, students may not freely express their opinions about institutions, courses, and lecturers.

Social Media

In teaching-learning improvement, Chauhan et al. (2019) illustrated the impact of aspects extracted from student's comments, collected from social media. Ortigosa et al. (2014) presented a SA model that combines lexical-based and machine learning techniques to extract information about the student's sentiments from Facebook messages in Spanish. Altrabsheh et al. (2013) collected students' tweets in English about their courses in real-time and used SA to improve teaching by some machine learning techniques. Altrabsheh et al. (2013), Sivakumar and Reddy (2017), and Nasim et al. (2017) also performed SA in English on the students' tweets to understand the students' opinions. Similarly, Kamisli Ozturk et al. (2017) have fetched students' tweets related to Anadolu University's open and distance education system and performed SA for Turkish to understand students' sentiments toward this institution. In another aspect, Kandhro et al. (2019) focused on teachers' evaluation and performed SA to identify the student sentiments from the piece of text. Also analyzing the student feedback on learning material, Cobos et al. (2019) have developed a tool to extract and analyze the opinion about their online courses. Gottipati et al. (2018) evaluated a solution using student feedback comments from seven undergraduate core courses taught at the School of Information Systems, Singapore Management University and compared rule-based methods and statistical classifiers to extract and summarize the explicit suggestions. To determine the current situation and trends of ODE students' use of Facebook, Firat et al. (2017) analyzed the density, centrality, and degrees of related Facebook groups and Facebook pages.

Nevertheless, as far as we know, the only SA study, which is about an open and distance education system in Turkish, is given by our sentiment research group (Kamisli Ozturk et al., 2017). Our previous study applied SA for a smaller dataset by using just one classifier as Naive Bayes. Here, our main contribution is the application of SA with different classifiers (Support Vector Machine, K-Nearest Neighborhood, Logistic Regression, Artificial Neural Networks) and vector space models (Bag of Words and Term Frequency-Inverse Document Frequency) to open and distance education system in Turkish.

SENTIMENT ANALYSIS

One of the main goals of SA is to exercise "Sentiment Polarity Classification" which is used to obtain the semantic polarity (positive, negative, or neutral) of a text (Saglam et al., 2016). There are three main approaches for this goal in SA; Lexicon-based Approach, Machine-learning-based Approach, and Hybrid Approach. Pandey et al. (2017) classified these approaches into sub-categories, as shown in Figure 1.

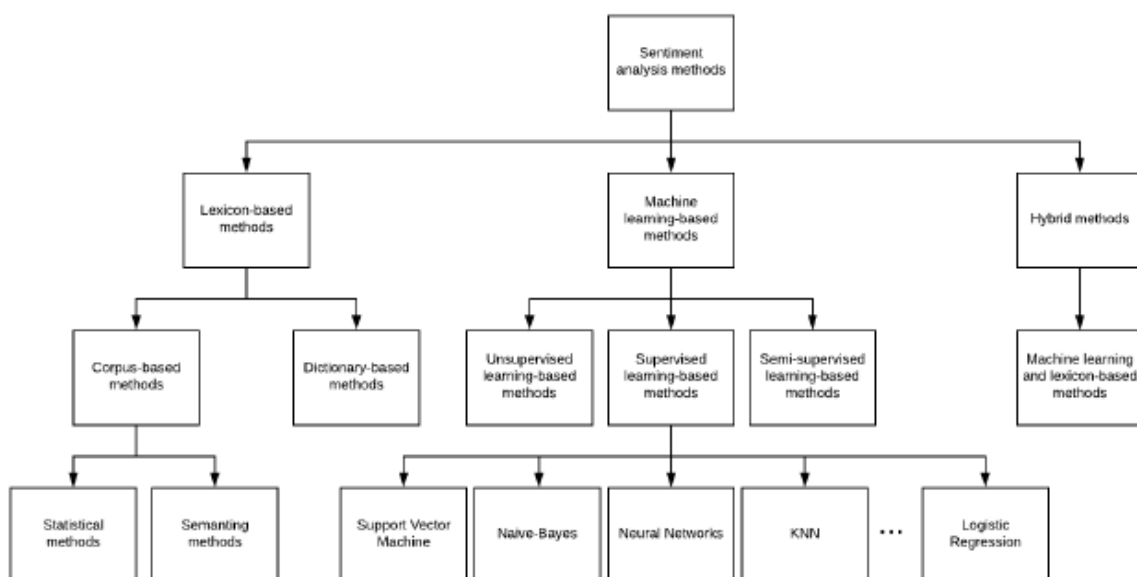


Figure 1. Classification of SA methods (Pandey et al., 2017)

The lexicon-based approach requires predefined polarity lexicons to obtain the sentiment polarity of the words in a text. There is no such dictionary in Turkish, while there are comprehensive predefined polarity lexicons in English, Dutch, Spanish, and Spanish. Turkish language’s distinctive characteristics, such as agglutinative, negation suffixes, make the Lexicon-based approach difficult; therefore, Lexicon-based SA studies in Turkish are based on translation-dependent lexicons or have relatively narrow focuses (Saglam et al., 2019). These studies are SentiTurkNet (Dehkharghani et al., 2016) by using Turkish WordNet (Bilgin et al., 2004) and SWNetTR (Ucan, 2014).

The machine learning-based approaches train a text classifier on the training set whose polarities are labeled by researchers. The first study in this category was performed by Pang et al. (Pang et al., 2002). They have shown that machine learning-based approaches classify with better performance than human-generated classifications. These approaches can be divided into three main categories: supervised, unsupervised, and semi-supervised learning-based methods. SVM, ANN, KNN, Naive Bayes, and LR have mostly used classifiers in supervised learning based category. In the Turkish language, the supervised learning based methods were used in many SA application areas such as politics (Kaya et al., 2012), marketing (Sternberg et al., 2018), (Karahoca et al., 2019)), and music (Durahim et al., 2018).

The last one, the hybrid approach, combines machine learning and lexicon-based approaches. There is only one study in literature until now (Ersahin et al., 2019), which applies the hybrid approach in Turkish.

In general, SA is performed at three levels: document, sentence, and aspect levels. The document level SA finds out the whole sentiment of a document, whereas the sentence level deals with each sentence individually. Aspect-level SA aims to classify sentiment according to specific characteristics (features) of an entity. Aspect-level SA can make different sentiment classifications for different characteristics of the same entity.

Machine learning-based methods are chosen in our study because there is no comprehensive polarity lexicon in Turkish, nor is there a narrow polarity lexicon containing terms related to ODE. Besides, (Turkmenoglu & Tantug, 2014), (Sevindi, 2013), (Ogul & Gonenc, 2016) showed that machine learning approaches perform better than lexicon-based in the Turkish language.

METHODOLOGY

The flowchart of the proposed methodology is given in Figure 2. In the following subsections, the steps of the flowchart will be explained in detail.

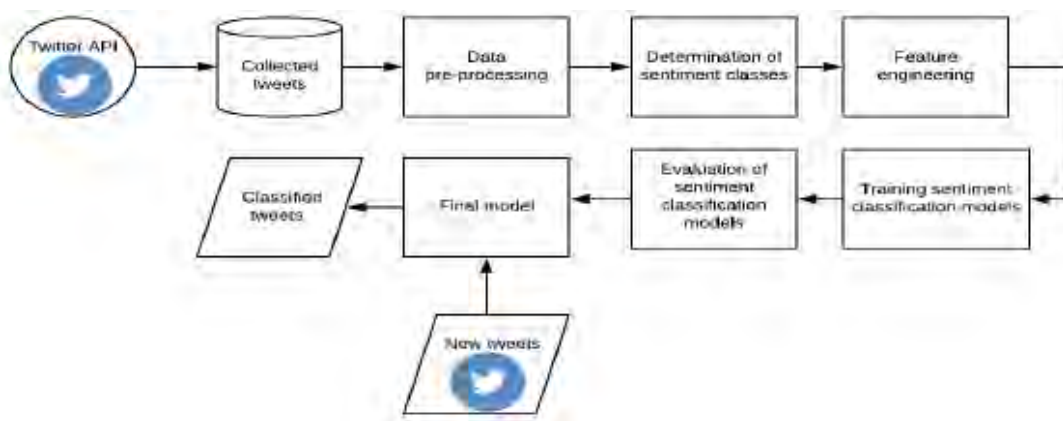


Figure 2. Flowchart of proposed methodology

Data Collection

Tweets were collected via Twitter API that returns public Tweets with a specific hashtag one week before the final examination and one week after the examination. We have used “aof”, “aof”, “acikogretim”, “acik ogretim”, “acikogretim”, “acik ogretim” hashtags, which are related to the ODE, to fetch the tweets from Twitter. At the end of this step, 63699 tweets were collected as a dataset.

Data Pre-processing

Data pre-processing involves four steps in our study; cleaning, normalization, tokenization, and stop word removal. In the cleaning step; duplicated tweets, unrelated links, URLs, advertisements, and news were removed from the dataset using regular expression. Foreign languaged tweets were investigated with “Language Detection API”, and we detected 103 foreign languages. These tweets were also removed from the dataset. 4652 tweets have remained in the dataset after these processes.

Normalization is the process of converting all text to the same case, eliminating punctuations, converting numbers to words, and so on. In the normalization step, we eliminated punctuations in the dataset and converted the dataset to a lower case.

Tokenization is the process of splitting the raw string into meaningful tokens (Hardeniya, 2015). In the tokenization step, tweets were split into words as a token by whitespaces.

Stop word removal is one of the most commonly used preprocessing steps in SA to reduce the vector space and enhance the classifier performance. Stop words have no significance in some of the NLP tasks like information retrieval and classification, which means these words are not very discriminative (Hardeniya, 2015). In the stop words removing step, we removed the stop words, which in the stop words list (Can et al., 2008) and the hashtags used in the data collection step from the tweets.

Determination of Sentiment Classes

To determine a sentiment class of each tweet, the tweets in the dataset were manually labeled as positive, negative, or neutral by three different native Turkish speakers via group decision-making.

Feature Engineering

Text needs to be represented as numerical feature vectors before applying machine learning algorithms. Bag of words (BoW), term frequency-inverse document frequency (TF-IDF), is used as vector space models in our study. In both models, the sequence of words is unigram or 1-gram, which means each word in the text represents a single word.

BoW model generates a feature vector that contains the counts of each unique word in the text without using semantics and order of the words. Figure 3 illustrates the logic of the BOW model.

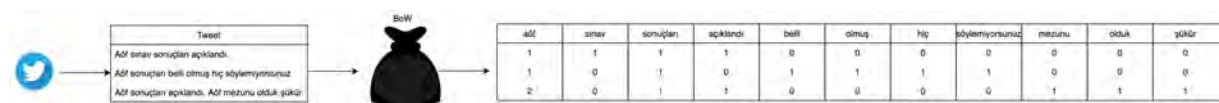


Figure 3. An example of bag-of-words representation

TF-IDF is a well-known method to evaluate a word’s importance in a text and multiplication of term frequency (TF) and the inverse document frequency (IDF). (TF) of a particular term (t) is calculated as the number of times a term occurs in a document to the total number of words in the text (Ahuja et al., 2019). IDF is the log of the inverse probability of a term being in the text.

Raw term frequency ($tf(t,d)$) is the number of times a term t occurs in a document (Raschka, 2019). The term frequency-inverse document frequency $tf-idf(t,d)$ is calculated as given in Equation (1).

$$tf-idf(t,d) = tf(t,d) \times idf(t,d) \quad (1)$$

Based on Equation (1) inverse document frequency $idf(t,d)$ can be calculated by Equation (2):

$$idf(t, d) = \log \frac{n_d}{1 + df(d, t)} \quad (2)$$

Sentiment Classification

To determine a sentiment class of each tweet, the tweets in the dataset were manually labeled as positive, negative, or neutral by three different native Turkish speakers via group decision-making. Although Naïve Bayes classifier (NBC) is often preferred in SA, the success of NBC in our previous study (Kamisli Ozturk et al., 2017) did not exceed 60%. Therefore, in this study, we used Support Vector Machine (SVM), Logistic Regression (LR), K-nearest neighbor (KNN), and Artificial Neural Network (ANN) to classify the tweets into positive, negative, or neutral sentiments.

SVM, which Vladimir N. Vapnik and Alexey Ya invented, is a widely used optimization-based supervised machine learning algorithm for classification and regression. The aim of the SVM is to find a hyperplane in N-dimensional space (N-the number of features) that correctly classifies the dataset by maximizing the margin between the two classes. Since we had a multiclass imbalanced dataset, we applied one-vs-one (OvO) strategy for handling multiclass classification for an imbalanced dataset in SVM. This strategy selects a pair of classes from a set of n classes and develops a binary classifier for each pair, so it requires to train classifiers. At prediction time, all classifiers are applied to the test dataset, and the class, which got the highest number of predictions, gets predicted.

Logistic Regression (LR) is a statistical classification model for the prediction of a binary categorical variable. The aim of Logistic Regression is to find a relationship between features and the probability of the outcome. To limit the probability of outcome value between 0 and 1, the sigmoid function is used for LR. Here, we applied one-vs-rest (OvR) strategy for handling multiclass classification. This strategy selects one class from a set of n classes and develops a binary classifier against all the other classes, so it requires training n classifiers.

KNN is a simple, lazy learner, similarity-based, and non-parametric algorithm used for regression and classification. The basic idea behind the KNN is finding the k samples in the training dataset that are most similar to the point that to be classified. The new data point's class label is then determined by a majority vote among its k nearest neighbors (Raschka, 2019). We used the Euclidean distance as a similarity metric in this study.

Inspired by biological systems, mainly by research into the human brain, ANNs are able to learn from and generalize from experience (Zhang et al., 1998). A basic ANN model consists of one input layer, one hidden layer, and one output layer. The processing elements, which are called nodes or neurons, are interconnected. An ANN model can be described as a directed graph in which each node i perform a transfer function f_i of the form.

$$y_i = f_i(\sum_{j=1}^n w_{ij}x_j - \theta_i) \quad (3)$$

Where y_i is the output of the node i , x_j is the j^{th} input to the node, w_{ij} is the connection weight between nodes i and j (Yao, 1999). θ_i is the bias term of each hidden node and transfer function f_i is a nonlinear function such as sigmoid and tanh.

Parameter Setting

Parameter setting is a crucial issue to improve the performance of classifiers. We used the grid search technique for finding the optimal combination of parameter values of these classifiers.

Evaluation Measures

The selection of the evaluation measure plays a vital role in understanding the classifier performance correctly. While selecting the evaluation measure, the ratio of each class in a dataset should be considered. If the dataset is imbalanced, which means the class distribution is not uniform among the classes, F-score can be used as an evaluation measure.

In this study, we have three sentiment classes as ‘neutral’, ‘positive’, and ‘negative’ and the sentiment class distribution of tweets in our data set is 1075 negative, 3345 neutral, and 232 positive. It is clearly seen that our dataset is imbalanced; therefore F-score, which is the harmonic mean of precision and recall, was used to compare the performance of the classifiers. To calculate the F-score, confusion matrix is constructed as shown in Table 1. Equation (4), (5), and (6) represent the computation of precision, recall, and F-score, respectively.

Table 1. Confusion matrix

		Actual Class	
		Yes	No
Predicted Class	Yes	True Positive (TP)	False Positive (FP)
	No	False Negative (FN)	True Negative (TN)

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$Recall = \frac{TP}{FN+TP} \quad (5)$$

$$F - score = 2 \frac{Precision+Recall}{Precision+Recall} \quad (6)$$

Computational Results

In order to generalize the models, we have used 10-fold cross-validation for reporting the results. Python is used to implement SA classification models. Tables 2 provides the F-score of classifiers according to TF-IDF and BoW model. It is observed that the LR classifier outperforms the other classifiers for both of the two vector space models. On the other hand, ANN and SVM give the F-score values that are not different from LR, and only KNN’s F-score is less than 64%. When we examine the results in terms of vector models, the BoW model gives better classification success rates for SVM, KNN, and ANN rather than the TF-IDF model. Imbalanced distributed sentiment polarities can explain the reason for this in the dataset. In the TF-IDF model, distinctive words for sentiment polarity determination took lower IDF values because of imbalanced distribution.

Table 2. F-score of classifiers according to TF-IDF and BoW model

	SVM	KNN	LR	ANN
BoW	0.7470	0.6341	0.7532	0.7411
TF-IDF	0.7350	0.6203	0.7584	0.7039

The LR and TF-IDF model’s detailed results, which give the best F-score, are shown in the graphs in Figure 4. Each graphic in this figure indicates the number of tweets and accurately predicted tweets according to the sentiment classes in the test dataset for each fold. Satisfactory results were obtained in the negative and

neutral emotion classes with the best model available. Since the number of tweets belonging to the positive classes was lower in the dataset than the other sentiment classes, the model had difficulty in learning the positive classes class.

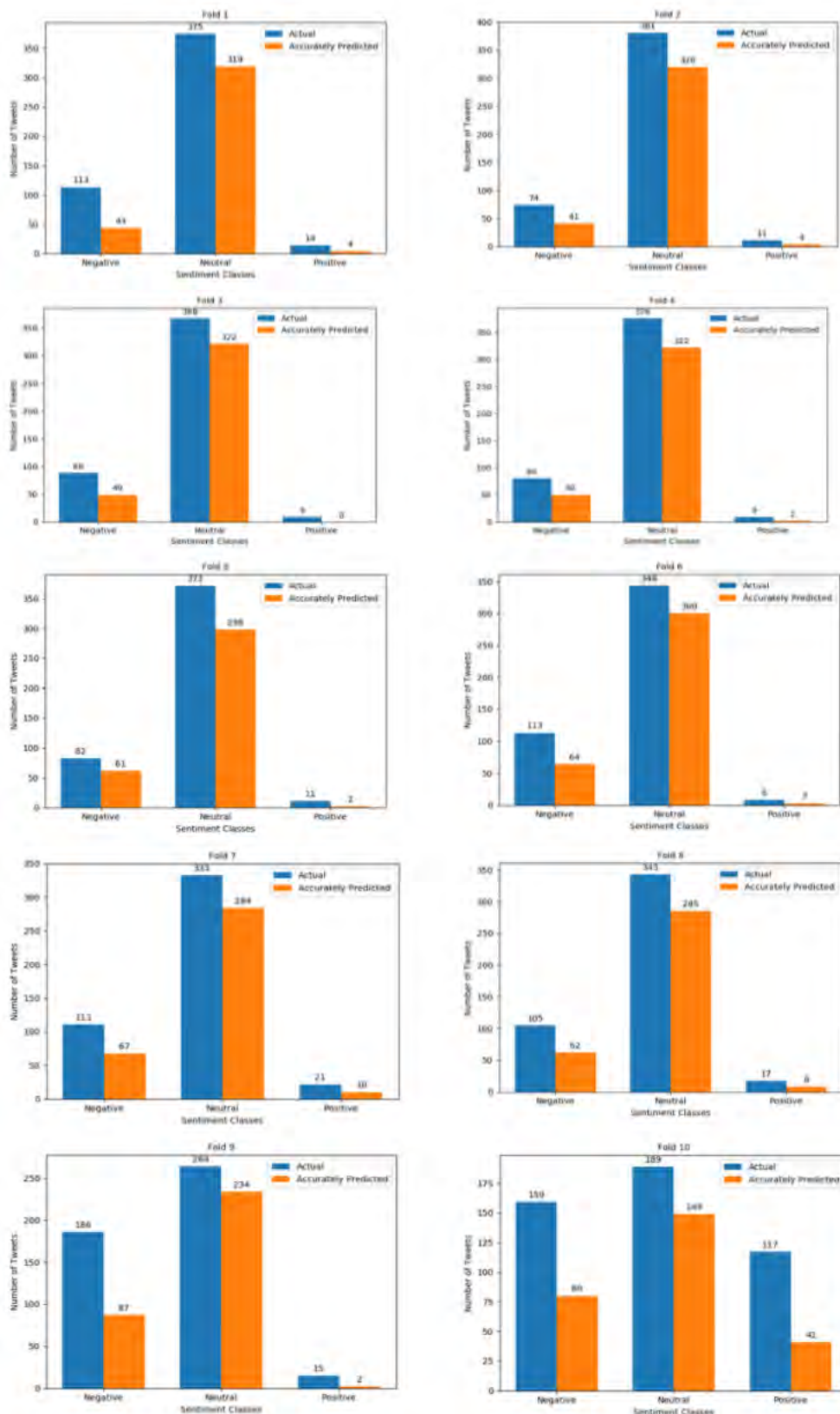


Figure 4. LR and TF_IDF model's detailed results for each fold

DISCUSSIONS

Learning analytics (Siemens, 2010) is the research field that combines techniques such as educational data mining, social networks analysis, sentiment analysis, and educational data conceptual modeling, to gather information and gain knowledge about the function and the results of an educational course at different levels. More specifically, through the learning analytics, an early detection of those students who need special help or are at risk of failure may occur, or some learning tools can be recommended, appropriate to the students' needs, or further suggestions for any decisions needed to be taken may be made (Kagklis et al., 2015). As given in Kagklis et.al. (2015), the use of social media and microblogging, virtual worlds, chat rooms, online discussion fora, blogs, and other web-based tools, synchronous or asynchronous, are emerging as useful and supportive tools to the educational process (Groves & O'Donoghue, 2009; Carsten et al., 2010). These tools can increase students' motivation and participation by making them able to determine the content of their discussions and define their educational needs (Choi et al., 2005; Bradley & McDonald, 2011).

Some studies use questionnaires to collect and analyze students' opinions about the ODE system (Aydin & Ozturk, 2017; Gurbuz, 2014; Kan Kilic et al., 2020). Generally, various evaluation questionnaires are applied to students and educators in order to conduct an internal evaluation in educational institutions. These surveys contain questions about faculty, courses, exams, and institutional infrastructure. However, with surveys' help, users' feelings and thoughts about the system may not be fully obtained. Based on the questionnaires' questions, the importance of the analysis of the texts is revealed regarding the subjective answers required for making decision-making data mining. Moreover, social media-based data can be collected faster and inexpensively than traditional questionnaires. Therefore, it is obviously important to benefit not only from surveys but also from social media platforms in which students express their feelings and thoughts. Based on these social media and SA aspects, in this study to address our first research question "How can we automatically classify the sentiment of the students' Turkish tweets about the ODE system for understanding the students' real sentiments about the system?", the flowchart of the proposed methodology was presented as a roadmap to classify the sentiments of the students' Turkish tweets about the ODE system.

With respect to our second research question, "Can different vector space models increase the success of sentiment classification?", using different space models gives different sentiment classification success as observed from Table 2. While SVM, KNN, and ANN gave better classification success with the BoW vector space model than the TF-IDF vector space model, LR gave better classification success with the TF-IDF vector space model than the BoW vector space.

In the literature review section, when the sentiment analysis for education-related studies is examined, it has been seen that there are not many studies related to ODE. Especially, when the sentiment analysis for ODE studies are considered, this study has a novelty for Turkish SA for ODE in Turkey. This study collected more social media data and used different vector space models and classification algorithms, and sentiment classification success increased from 56% (Kamisli Ozturk et al., 2017) to 75%. This result means that the students' opinions and sentiments can understand more accurately by the ODE administrators. This provides administrators with a decision-making environment that can benefit from distance education quality activities. Social media's important role in receiving feedback on the quality of education service provided by distance education institutions has been considered and emphasized.

Notwithstanding these significant results were obtained in the present study, this research has two limitations that should be mentioned. This study is limited by using a single data source, Twitter. Some students can share their opinions about the ODE system on social media platforms other than Twitter. Also, the lack of a comprehensive polarity lexicon in Turkish limits the methods used in the study. Lexicon-based SA cannot be performed in this study because of this limitation.

Besides SA studies, there are a limited number of studies using social media data about ODE in Turkey. One of these studies is presented by Firat et al. (2017) to use holistic social network analysis on Facebook to reveal social media use trends in ODE in Turkey. They focused on Facebook groups and pages about the ODE and used like counts of these groups/pages for social network analysis.

CONCLUSIONS

Good decisions about ODE require good information and datasets. Since Twitter is a platform where people can freely share their thoughts, with this study, it was provided to reveal the real opinions of students about the open and distance education system. With the application of SA to students' tweets about the open and distance education system, it is aimed to reveal the students' perceptions, opinions, and complaints about the ODE system using machine learning algorithms. Throughout this work, we presented several SA models to extract an ODE institution's students' opinions in Turkey. The most critical problem we encountered in this study is the lack of a comprehensive polarity lexicon in Turkish. To overcome this problem, sentence-level SA with BoW and TF-IDF vector models and SVM, KNN, LR, and ANN machine learning algorithms were considered. As reported in the computational results, the LR classifier outperformed the other classifiers for both of the two vector space models in F-scores for the considered dataset.

Considering the comparison of different SA models can be a guiding study for open and distance institutions. As it is given in the literature, SA studies in the Turkish language have not been considered in the literature intensively. From this point of view, this study will present a roadmap for researchers to study SA in Turkish. Our findings are encouraging researchers to focus on different vector space models for imbalanced textual data on the technical contribution side.

As a result of the presented SA models, ODE managers can develop strategies that will increase student satisfaction, based on the tweets whose feelings for the Open and Distance Education System are negative and positive. Based on the positive sentiments, strategies can be developed to reinforce ODE's positive transactions and ensure its sustainability. Besides, negative sentiments can be used in decision making, allowing an understanding of student dissatisfaction and how the institutions adopt actions to improve the quality of teaching and examination processes used. The presented SA models within the scope of the study can be utilized at any time. Thus, the reflection of managerially changing conditions to the system users can be continuously monitored.

As future work, we are planning to develop a comprehensive Turkish lexicon for performing lexicon-based SA. We are also planning to diversify our data sources by collecting data from different social media platforms such as Facebook and Instagram. In this manner, limitations of the presented work can overcome with these strategies.

Authors' Note: This study is supported by Anadolu University Scientific Research Projects Committee (AUBAP-1604E173) and Eskisehir Technical University Scientific Research Projects Committee (ESTUBAP-19ADP048).

BIODATA and CONTACT ADDRESSES of AUTHORS



Zeliha ERGUL AYDIN is a full-time research assistant and a Ph.D. student at Eskisehir Technical University Industrial Engineering Department. Her research interests include machine learning, data analytics, text mining, metaheuristics, operations research and optimization. She earned a BS and an MS degree in Industrial Engineering from Anadolu University. She worked as a business analyst for 3 years in the IT team which preparing Anadolu University Open and Distance Education System examinations. She has 4 journal articles in international indexes and 10 studies presented at international meetings.

Zeliha ERGUL AYDIN
Department of Industrial Engineering, Faculty of Engineering
Address: Eskisehir Technical University, 26555, Eskisehir, Turkey
Phone: +90 222 321 35 50
E-mail: zergul@eskisehir.edu.tr



Assoc. Prof. Dr. Zehra KAMISLI OZTURK is a faculty member at Eskisehir Technical University Industrial Engineering Department. She completed her doctorate at Eskisehir Osmangazi University. She has studied at Ballarat University as an academic visitor during her Ph.D. She worked as a research assistant at Anadolu University Open Education Faculty between 2002-2012. During this period, she took part in system analysis, simulation, optimization and data analytics projects related to distance education system. Her research interests include operations research, optimization, multi-criteria decision making, and data analytics. She has given many courses on operations research, multi-criteria decision making, and simulation; she has done master's and doctoral thesis advisors; participated in national and international projects related to her field as both researcher and manager.

Zehra KAMISLI OZTURK

Department of Industrial Engineering, Faculty of Engineering
Address: Eskisehir Technical University, 26555, Eskisehir, Turkey
Phone: +90 222 321 35 50
E-mail: zkamisli@eskisehir.edu.tr



Dr. Zeynep Idil ERZURUM CICEK is a full-time research assistant at Eskisehir Technical University Industrial Engineering Department. Her research interests include machine learning, text mining, metaheuristics, operations research and optimization. She earned a BS and an MS degree in Industrial Engineering from Anadolu University and a Ph.D. degree in Industrial Engineering from Eskisehir Technical University. She worked as a junior production-planning specialist at Eczacibasi Building Products and business analyst in Anadolu University Computer Research Center. She has 4 journal articles in international indexes, 3 book chapters and 10 studies presented at international meetings.

Zeynep Idil ERZURUM CICEK

Department of Industrial Engineering, Faculty of Engineering
Address: Eskisehir Technical University, 26555, Eskisehir, Turkey
Phone: +90 222 321 35 50
E-mail: zierzurum@eskisehir.edu.tr

REFERENCES

- Ahmed, S., Jaidka, K., & Cho, J. (2016). The 2014 Indian elections on Twitter: A comparison of campaign strategies of political parties. *Telematics and Informatics*, 33 (4), 1071-1087.
- Ahuja, R., Chug, A., Kohli, S., Gupta, S., & Ahuja, P. (2019). The Impact of Features Extraction on the Sentiment Analysis. *Procedia Computer Science*, 152, 341-348.
- Ali, M., & Abdel-Haq, M. K. Bibliographical Analysis of Artificial Intelligence Learning in Higher Education: Is the Role of the Human Educator and Educated a Thing of the Past?. In *Fostering Communication and Learning With Underutilized Technologies in Higher Education* (pp. 36-52). IGI Global.)
- Altrabsheh, N., Gaber, M., & Cocea, M. (2013). SA-E: Sentiment Analysis for Education. *International Conference on Intelligent Decision Technologies*, 255, pp. 353-362.
- Aung, K., & Myo, N. (2017). Sentiment Analysis of Students' Comment Using Lexicon Based Approach. *Computer and Information Science (ACIS)*, (pp. 149-154). Wuhan.
- Aydin, S., & Ozturk, A. (2017). Anadolu Universitesi Acikogretim Sistemindeki ogrenenlerin sinav ve buro hizmetlerine yonelik goruslerinin degerlendirilmesi. *Acikogretim Uygulamalari ve Arastirmalari Dergisi*, 3 (3), 64-79

- Bansal, B., & Srivastava, S. (2018). On predicting elections with hybrid topic based sentiment analysis of tweets. *Procedia Computer Science*, 135, 346-353.
- Barbier, G., & Liu, H. (2011). Data Mining in Social Media. In C. Aggarwal, *Social Network Data Analytics* (pp. 327-352). Springer, Boston, MA.
- Bilgin, O., Cetinoglu, O., & Oflazer, K. (2004). Building a wordnet for Turkish. *Romanian Journal of Information Science and Technology*, 7 (1-2), 163-172.
- Bradley, A.J., & McDonald, M.P. (2011, October 26). *Social Media versus Knowledge Management [Blog post]*. *Harvard Business Review Blog*. Retrieved 10-1-2014 from <http://blogs.hbr.org/2011/10/social-media-versus-knowledge/>
- Can, F., Kocberber, S., Balcik, E., Kaynak, C., Ocalan, H., & Vursavas, O. (2008). Information Retrieval on Turkish Texts. *Journal of the American Society for Information Science and Technology*, 59 (3), 407-421.
- Carsten, U., Boreau, K., & Stepanyan, K. (2010). Who students interact with? A social network analysis perspective on the use of Twitter in Language Learning. In M. Wolpers, P. Kirschner, M. Scheffel & V. Dimitrova (Eds.), *Proceedings of 5th European Conference on Technology Enhanced Learning* (pp. 432-437). Sankt Augustin: Springer.
- Chauhan, G., Agrawal, P., & Meena, Y. (2019). Aspect-Based Sentiment Analysis of Students' Feedback to Improve Teaching–Learning Process. *Information and Communication Technology for Intelligent Systems* (pp. 259-266). Singapore: Springer.
- Cobos, R., Jurado, F., & Villen, A. (2019). Moods in MOOCs: Analyzing Emotions in the Content of Online Courses with edX-CAS. *2019 IEEE Global Engineering Education Conference (EDUCON)*, (pp. 1467-1474). Dubai.
- Choi, I., Land, S. M., & Turgeon, A. J. (2005). Scaffolding peer-questioning strategies to facilitate metacognition during online small group discussion. *Instructional Science*, 33, 483-511.
- Dehkharghani, R., Saygin, Y., & Yanikoglu, B. (2016). SentiTurkNet: a Turkish polarity lexicon for sentiment analysis. *Lang Resources & Evaluation*, 50, 667–685.
- de Oliveira, A.B., Alves, A.L.F., & de Souza Baptista, C. (2021). Using opinion mining in student assessments to improve teaching quality in universities. In *Intelligent Systems Design and Applications: 19th International Conference on Intelligent Systems Design and Applications (ISDA 2019)* Held December 3-5, 2019 (Vol. 1181, p. 225). Springer Nature.
- Dolianiti, F., Iakovakis, D., Dias, S., Hadjileontiadou, S., Diniz, J., & Hadjileontiadis, L. (2019). Sentiment Analysis Techniques and Applications in Education: A Survey. *International Conference on Technology and Innovation in Learning, Teaching and Education* (pp. 412-427). Cham: Springer.
- Dragoni, M., Poria, S., & Cambria, E. (2018). OntoSenticNet: A Commonsense Ontology for Sentiment Analysis. *IEEE Intelligent Systems*, 33 (3), 77-85.
- Durahim, A., Coskun Setirek, A., Basarir Ozel, B., & Kebapci, H. (2018). Music emotion classification for Turkish songs using lyrics. *Pamukkale University Journal of Engineering Sciences*, 24 (2), 292-301.
- Ersahin, B., Aktas, O., Kilinc, D., & Ersahin, M. (2019). A hybrid sentiment analysis method for Turkish. *Turkish Journal of Electrical Engineering & Computer Sciences*, 27, 1780-1793.
- Firat, M., Altinpulluk, H., Kilinc, H., & Buyuk, K. (2017). Determining open education related social media usage trends in Turkey using a holistic social network analysis. *Educational Sciences: Theory & Practice*, 17(4), 1361-1382.
- Georgoula, I., Pournarakis, D., Bilanakos, C., Sotiropoulos, D., & Giaglis, G. (2015). Using Time-Series and Sentiment Analysis to Detect the Determinants of Bitcoin Prices. *Ninth Mediterranean Conference on Information Systems (MCIS)*. Samos.
- Goloshchapova, I., & Andreev, M. (2017). Measuring inflation expectations of the Russian population with the help of machine learning. *VOPROSY ECONOMIKI*, 6, 71-93.
- Gottipati, S., Shankararaman, V., & Lin, J. (2018). Text analytics approach to extract course improvement suggestions from students' feedback. *RPTTEL*, 13 (6).

- Groves, M., & O'Donoghue, J. (2009). Reflections of Students in Their Use of Asynchronous Online Seminars. *Educational Technology & Society*, 12(3), 143-149.
- Gurbuz, F. (2014). Students' Views On Distance Learning In Turkey: An Example Of Anadolu University Open Education Faculty . *Turkish Online Journal of Distance Education* , 15 (2) , 239-250
- Hardeniya, N. (2015). NLTK Essentials. Packt Publishing.
- Jaidka, K., Ahmed, S., Skoric, M., & Hilbert, M. (2019). Predicting elections from social media: a three-country, three-method comparative study. *Asian Journal of Communication*, 29 (3), 252-273.
- Kagklis, V., Karatrantou, A., Tantoula, M., Panagiotakopoulos, C. T., & Verykios, V. S. (2015). A Learning Analytics Methodology for Detecting Sentiment in Student Fora: A Case Study in Distance Education, *European Journal of Open, Distance and E-Learning*, 18(2), 74-94.
- Kamisli Ozturk, Z. (2012). Operations Research Applications in Higher Open and Distance Learning Systems. In G. Bell, J. Warwick, & P. Galbraith, *Higher Education Management and Operational Research* (pp. 253-268). Rotterdam: SensePublishers.
- Kamisli Ozturk, Z., Erzurum Cicek, Z., & Ergul, Z. (2017). Sentiment analysis: An application to Anadolu University. *Acta Physica Polonica A*, 132 (3), 753-755.
- Kan Kilinc, B , Yazici, B , Gunsoy, B , Gunsoy, G . (2020). Perceptions and Opinions of Graduates about the Effects of Open and Distance Learning in Turkey. *Turkish Online Journal of Distance Education*, 21 (1) , 121-132
- Kandhro, I., Wasi, S., Kumar, K., Rind, M., & Ameen, M. (2019). Sentiment Analysis of Students' Comment using Long-Short Term Model. *Indian Journal of Science and Technology*, 12 (8), .
- Karahoca, A., Karahoca, D., & Evirgen, E. (2019). SENTIMENT ANALYSIS OF TURKISH TWEETS BY DATA MINING METHODS. *International Journal of Mechanical Engineering and Technology (IJMET)*, 10 (01), 915-925.
- Kaya, M., Fidan, G., & Toroslu, I. (2012). Sentiment Analysis of Turkish Political News. *2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology*, (pp. 174-180). Macau.
- Kechaou, Z., Ammar, M., & Alimi, A. (Amman). Improving e-learning with sentiment analysis of users' opinions. *2011 IEEE Global Engineering Education Conference (EDUCON) – "Learning Environments and Ecosystems in Engineering Education"* (pp. 1032-1038). IEEE.
- Kim, S., & Calvo, R. (2010). *Sentiment analysis in student experiences of learning*. EDM. Pittsburgh.
- Kušen, E., & Strembeck, M. (2018). Politics, sentiments, and misinformation: An analysis of the Twitter discussion on the 2016 Austrian Presidential Elections. *Online Social Networks and Media*, 5, 37-50.
- Liu, B. (2012). Sentiment Analysis and Opinion Mining. *Synthesis Lectures on Human Language Technologies*, 5 (1), 1-167.
- Nasim, Z., Rajput, Q., & Haider, S. (2017). Sentiment analysis of student feedback using machine learning and lexicon based approaches. *2017 International Conference on Research and Innovation in Information Systems (ICRIIS)*, (pp. 1-6). Langkawi.
- Nguyen, Q., Brunisholz, K., Yu, W., McCullough, M., Hanson, H., Litchman, M., et al. (2017). Twitter-derived neighborhood characteristics associated with obesity and diabetes. *Scientific Reports*, 1-10.
- Nguyen, T., Shirai, K., & Velcin, J. (2015). Sentiment analysis on social media for stock movement prediction. *Expert Systems with Applications*, 42 (24), 9603-9611.
- Ogul, B., & Gonenc, E. (2016). Sentiment classification on Turkish hotel reviews. *2016 24th Signal Processing and Communication Application Conference (SIU)*, (pp. 497-500). Zonguldak.
- Opuszko, M., Bode, L., & Ulbricht, S. (2018). Analysis of Twitter Communication During the 2017 German Federal Election. *DATA ANALYTICS 2018 : The Seventh International Conference on Data Analytics*. Athens.
- Ortigosa, A., Martín, J., & Carro, R. (2014). Sentiment analysis in Facebook and its application to e-learning. *Computers in Human Behavior*, 31, 527-541.

- Oscar, N., Fox, P., Croucher, R., Wernick, R., Keune, J., & Hooker, K. (2017). Machine Learning, Sentiment Analysis, and Tweets: An Examination of Alzheimer's Disease Stigma on Twitter. *The Journals of Gerontology: Series B*, 72 (5), 742-751.
- Pandey, A., Rajpoot, D., & Saraswat, M. (2017). Twitter sentiment analysis using hybrid cuckoo search method. *Information Processing and Management*, 53 (4), 764-779.
- Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment Classification using Machine Learning Techniques. *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2002* (pp. 79-86). Association for Computational Linguistics.
- Pozzi, F., Fersini, E., Messina, E., & Liu, B. (2016). Challenges of Sentiment Analysis in Social Networks: An Overview. In *Sentiment Analysis in Social Networks* (pp. 1-11). Morgan Kaufmann.
- Raschka, S. (2019). *Python Machine Learning*. Packt Publishing.
- Saglam, F., Genc, B., & Sever, H. (2019). Extending a sentiment lexicon with synonym-antonym datasets: SWNetTR++. *Turkish Journal of Electrical Engineering & Computer Sciences*, 27, 1806-1820.
- Saglam, F., Sever, H., & Genc, B. (2016). Developing Turkish sentiment lexicon for sentiment analysis using online news media. *IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA)*, 2016 (pp. 1-5). Agadir.
- Sangam, S., & Shinde, S. (2019). A Novel Feature Selection Method Based on Genetic Algorithm for Opinion Mining of Social Media Reviews. In Information, Communication and Computing Technology. ICICCT 2018. *Communications in Computer and Information Science*, (Vol. 835). Singapore: Springer.
- Sevindi, B. (2013). Turkece Metinlerde Denetimli Ve Sozluk Tabanlı Duygu Analizi Yaklaşımlarının Karsılaştırılması. Gazi Üniversitesi, Fen Bilimleri Enstitüsü.
- Siemens, G. (2012). Learning analytics: envisioning a research discipline and a domain of practice. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, ACM, New York, NY, USA (2012), pp. 4-8
- Sivakumar, M., & Reddy, U. S. (2017). Aspect based sentiment analysis of students opinion using machine learning techniques. *2017 International Conference on Inventive Computing and Informatics (ICICI)*, (pp. 726-731). Coimbatore.
- Song, D., Lin, H., & Yang, Z. (2017). Opinion Mining in e-Learning System. *2007 IFIP International Conference on Network and Parallel Computing*, (pp. 788-792).
- Sorour, A., Atkins, A. S., Stanier, C. F., & Alharbi, F. D. (2020, February). The Role of Business Intelligence and Analytics in Higher Education Quality: A Proposed Architecture. In *2019 International Conference on Advances in the Emerging Computing Technologies (AECT)* (pp. 1-6). IEEE.
- Sternberg, F., Pedersen, K., Ryelund, N., Mukkamala, R., & Vatrappu, R. Analysing Customer Engagement of Turkish Airlines Using Big Social Data. *2018 IEEE International Congress on Big Data*, (pp. 74-81).
- Thomas, E., & Galambos, N. (2004). What Satisfies Students? Mining Student-Opinion Data with Regression and Decision Tree Analysis. *Research in Higher Education*, 45 (3), 251-269.
- Turkmenoglu, C., & Tantug, A. (2014). Sentiment Analysis in Turkish Media. In Workshop on Issues of Sentiment Discovery and Opinion Mining, *International Conference on Machine Learning (ICML)*.
- Ucan, A. (2014). Automatic sentiment dictionary translation and using in sentiment analysis. Hacettepe University.
- Ullah, M. (2016). Sentiment Analysis of Student Feedback: A Study towards Optimal Tools. *2016 International Workshop on Computational Intelligence (IWCI)*, (pp. 175-180). Dhaka.
- Wen, M., Yang, D., & Rose, C. (2014). Sentiment Analysis in MOOC Discussion Forums: What does it tell us? *Proceedings of the 7th International Conference on Educational Data Mining (EDM 2014)*.
- Yao, X. (1999). Evolving Artificial Neural Networks. *Proceedings of The IEEE*, 87 (9), 1423-1447.
- Zhang, G., Patuwo, B., & Hu, M. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, 14 (1), 35-62.