

Twitter Sentiment Analysis Based on Daily Covid-19 Table in Turkey

 Buket Kaya¹,  Abdullah Günay²

¹Corresponding Author; Firat University, Department of Electronics and Automation; bkaya@firat.edu.tr;
²Siirt University, Department of Call Center Services; abdullah.gunay@siirt.edu.tr;

Received 4 May 2021; Revised 26 October 2021; Accepted 3 November 2021; Published online 31 December 2021

Abstract

The coronavirus pandemic, which began to affect the whole world in early 2020, has become the most talked about agenda item by individuals. Individuals announce their feelings and thoughts through various communication channels and receive news from what is happening around them. One of the most important channels of communication is Twitter. Individuals express their feelings and thoughts by interacting with the tweets posted. This study aims to analyze the emotions of the comments made under the "daily coronavirus table" shared by the Republic of Turkey Ministry of Health and to measure their relationship with the daily number of cases and deaths. In the study, emotional classification of tweets was implemented using LSTM, GRU and BERT methods from deep learning algorithms. The results of all three algorithms were compared with the daily number of cases and deaths.

Keywords: Sentiment Analysis, Twitter, Deep Learning, NLP, LSTM, BERT, GRU, Covid19

1. Introduction

The rapid development of technology and related internet has created changes in people's social life. The fact that people want to share their activities, ideas and likes has made social media a popular platform and also an essential source of information. Today, while more than 4.5 billion people can surf the internet, almost 85% of them (3.8 billion) actively use social media. Twitter is still among the most effective social platforms, although its usage has dropped rapidly. As of 2020, the number of Twitter users is around 340 million. According to the Digital 2020 Global Outlook Report, Turkey ranks 6th globally with 11.8 million active Twitter users and 2nd in Europe [1].

With the Covid-19 virus outbreak that emerged in early 2020, people and governments found themselves in a crisis for which they were unprepared. With the pandemic, people began to physically stay away from social spaces. This situation has caused a change in individuals' behavior and daily habits. According to the study conducted by Ipsos [2], it has been revealed that individuals who follow the popular phrase 'Stay at Home' rule of the pandemic have acquired habits such as spending time in the kitchen, acquiring a habit of doing sports at home, taking advantage of online courses and trainings, and conducting video interviews. Later, these activities were shared on social media platforms. Examining the comments, sharing and discussions on social media platforms in order to observe the effects of the Covid-19 pandemic on the emotional world of people constitute an important resource for text mining. In its simplest definition, text mining is a type of data mining that sees what is written as a data source and analyzes this data. According to current data, more than 8,000 Tweets are shared on Twitter every second [3]. While writing the tweets, the words that were written inaccurately and without obeying the spelling rules, not being used in daily life and concepts specific to the social media environment have developed a unique language on Twitter. It is difficult to understand and analyze these tweets with human perception; Therefore, these data must be filtered, parsed and processed with natural language processing methods. There are many studies on the processing of data obtained from Twitter in the data mining sector by subjecting it to natural language processing One of them aims to predict pandemics with tweets posted between May and December 2009 [4]. Another study analyzed the side effects of drugs from tweets posted between the same year and months [5]. With 85% prediction success, social media dynamics changed human perception over time [6]. There is a study that analyzes perception over millions of tweets that tourists have sent for a touristic destination [7]. In another article, positive / negative classification was made with machine learning techniques over the comments shared on the

IMDB movie scaling platform [8]. Finally, there is a study on the analysis of emotions and thoughts in online environments [9].

Text mining and sentiment analysis aims to reveal the opinions, feelings and thoughts hidden in the comments and opinions shared by people on social media. Sentiment analysis has recently emerged as an active research analyzing people's opinions, feelings, and evaluations [10]. Working areas for sentiment analysis are Twitter and other social networks, blogs, forums and comments. Sentiment analysis studies conducted so far can be generally examined in two groups as dictionary and statistical (machine learning). The dictionary-based approach makes use of pre-prepared dictionaries of emotion concepts during classification. In the statistical / machine learning-based approach, machine learning algorithms and linguistic features are used during classification [11].

The purpose of this study is to automatically determine the emotionally positive / negative / neutral status of tweets posted by users on Twitter during the Covid-19 pandemic process. This study will be useful and important for decision makers by analyzing the mood of the society in the situation of social disasters and pandemics.

2. Related Work

Sabuncu and Atmış [12] recorded English tweets about Turkish Airlines with the R programming language and classified and interpreted the tweets using emotion analysis algorithm. Ataman and Ozguner [13] analyzed the tweets related to the Black Friday week in 2018 and analyzed the emotions contained in the tweets using the Python language and the Cognitus API. Kilimci [14] aimed to predict the direction of the Bist100 index with financial sentiment analysis. For this purpose, Kilimci has developed deep community models. In addition, in the system he developed, he achieved a high classification success of around 78% in Turkish and English data sets. In his study, [15] analyzed the positive, neutral and negative state of the shares after the social media platform Twitter subjected the shares of people to various pre-processes [11], on the other hand, compared the classification success using the multi-layer sensor (MLP), Naive Bayes (NB), Support Vector Machines (DVM) and Logistic Regression (LR) algorithms on user comments on an online book sales site. Chandra et al. [16] tried to find out whether tweets on Twitter are Islamophobic using sentiment analysis. The paper presented the CoronaBias dataset focusing on anti-Muslim hate, with more than 410,990 tweets from 244,229 users. This data set was used to perform longitudinal analysis. The study measured qualitative changes in the context associated with the Muslim community and performed macro and micro subject analysis to find common themes. Jianqiang et al. [17] presented a method of word embedding on large tweets based on unsupervised learning. The method uses semantic relationships between words in tweets and statistical features that occur together. Concerning the sentiment analysis, recently, Mungen et al. conducted a study on finding relationships between news and people's emotions in the Covid-19 pandemic [18].

Samuel et al. [19], using Tweets and R statistics software specific to coronavirus, determined the public sensitivity associated with the pandemic by emotion analysis method. The study achieved a 91% success rate for short tweets using the naive bayes-based classification method. In shorter tweets, it was observed that the logistic regression-based method remained at a success rate of 74%. On the other hand, it was determined that both methods performed relatively poorly in longer tweets. Chen et al. [20] used BERT to integrate data from emotion dictionaries. They showed that their approach consistently outperformed the state-of-the-art methods across all datasets. Lu et al. [21] first process the data they obtain through the BERT model. They then simulated linguistic functions in the sentence by combining Chinese grammar rules in the form of constraints with the Bi-gated recurrent neural network (GRU) and standardizing the output of adjacent positions. Wang et al. [22] analyzed 999,978 Weibo posts related to randomly selected COVID-19 from January 1, 2020 to February 18, 2020 in Sina Weibo, a popular Chinese social media. While the BERT model is used to classify sentiment categories, the TF-IDF model is more preferred to summarize the topics of the posts. Li et al. [23] presented a new method called GBCN, which uses a gate mechanism with context-sensitive direction placements to enhance and control BERT representation for perspective-based sentiment analysis. In the model created, SentiHood and SemEval-2014 data sets reached 88.0 and 92.9 F1 test results, respectively.

3. Data Description

3.1. Data Collection

Within the scope of the study, the emotions of individuals related to the COVID-19 pandemic, which is effective all over the world, were tried to be determined. It is planned to receive data from Twitter, a social media platform with a high rate of use in the world and in Turkey. However, due to the fact that there are a lot of tweets about COVID-19 and it is difficult to determine whether the tweets are related to the cases and deaths in Turkey or the world in general, a certain restriction was needed. At the same time, the fact that there are certain restrictions in terms of extracting data from Twitter required the narrowing of the area where the data will be collected. In this context, the comments made under the daily coronavirus tables announced by the Minister of Health of the Republic of Turkey on their official Twitter account were analyzed. Tweets about Turkey's daily coronavirus table announced by the Minister of Health are quoted, answered and retweeted by twitter users. While Figure 1 shows Turkey Daily Coronavirus Table, interaction of tweets explaining the Daily Coronavirus Table are given in Figure 2.



Figure 1 Turkey Daily Coronavirus Table

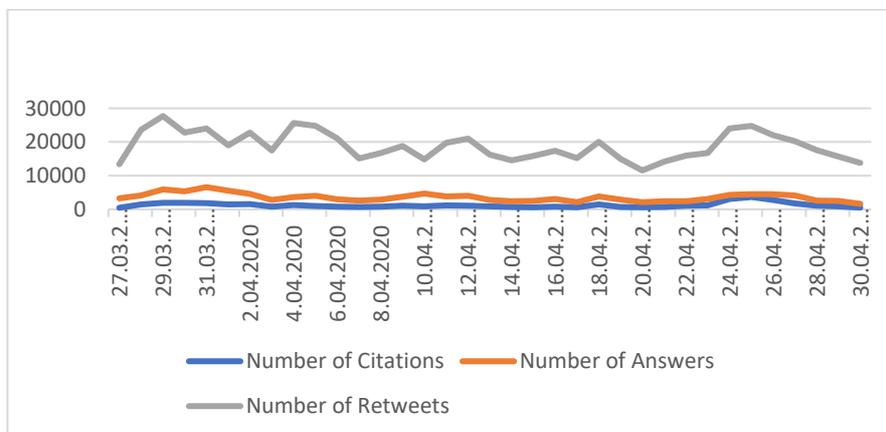


Figure 2 Interaction of Tweets Explaining the Daily Coronavirus Table

A total of 16240 tweets that were commented on the daily tables from March 27, 2020, when the coronavirus table was first announced in Turkey to April 30, 2020 using Twitter Premium API, were obtained through the Python program.

3.2. Data Pre-Processing

The data obtained from Twitter were subjected to various pre-processing using the Knime platform. Various filters are needed to process the words in the texts. In this context, repetitive expressions that are expressed as special characters, numbers and stop words in the texts and do not have any meaning on their own were found and filtered and all words were converted to lowercase letters. Then, using the Zemberek Library, which was prepared to process Turkish texts, word roots were found and separated into their elements. In addition, the usage frequencies of the words grouped in the "Bag of Words Creator" stage are calculated. The steps for preprocessing the data are given in Figure 3.

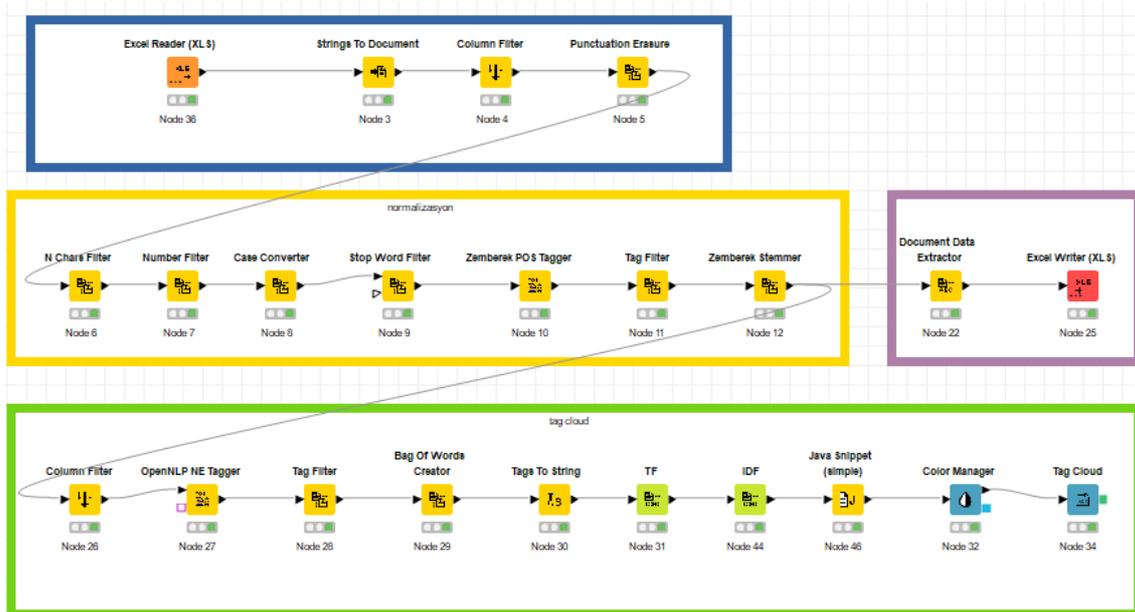


Figure 3 Pre-Processing of Data

The TF-IDF method is an approach that finds the relative frequency of words in a given document with the inverse ratio of words on the entire document corpus. The method uses two elements to determine the value: the i term frequency in the TF - j document and the reverse document frequency of the IDF - i term. TF-IDF can be calculated as follows [24]:

$$a_{ij} = tf_{ij}idf_i = tf_{ij} * \log_2\left(\frac{N}{df_i}\right) \quad (1)$$

In the formula, a_{ij} represents the weight of term i in document j . While tf_{ij} shows the frequency of the term i in document j , df_i gives to the document frequency of the term i in the collection. N is the number of documents in the collection.

As seen in Figure 4, using Zemberek-NLP, a natural language processing tool developed for Turkish, the word types in the data set were determined and the usage frequencies of the words were calculated. Figure 5 denotes word cloud derived from this data.

language. Many NLP implementations are based on language models that define probability distribution over words, characters or byte sequences in a natural language [25].

LSTM is a special type of RNN developed to address the vanishing / exploding problem encountered in RNNs (Figure 6). Similar to other RNNs, the output of the model is formed according to the input from the current time step and the output from the previous time step, and the current output is sent to the next time step. A memory cell (c_t) that stores the LSTM units state in random time intervals consists of three non-linear gates. These are an input gate (i_t), a forget gate (f_t), and output gate (o_t). The purpose of these gates is to regulate the flow of information entering and leaving the memory cell [26].

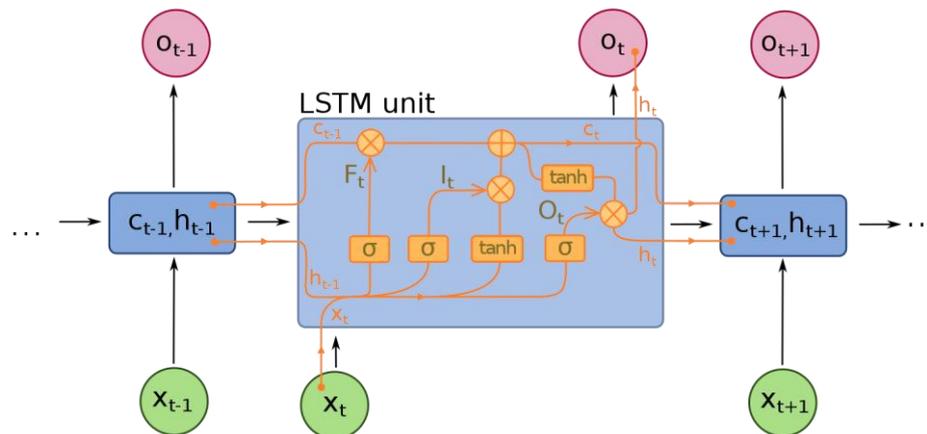


Figure 6 LSTM Model

GRU is an LSTM featured model that optimizes the LSTM network structure. When compared to the LSTM network structure, it is seen that the GRU network has two gateways, namely the update gate and the reset gate. This is because GRU is able to handle the long range long delay series prediction problem. While the reset gate determines the degree to which the information of the previous moment is ignored, the purpose of the update gate is to find out how close the information of the previous moment is to the current moment. A GRU model is shown in Figure 3. GRU has been noted to outperform LSTM units in terms of both CPU time convergence and parameter updates and generalization [23].

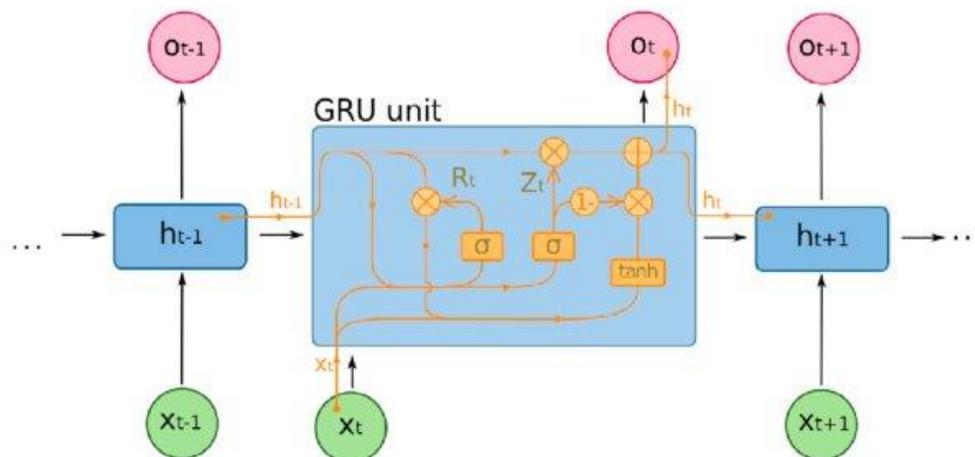


Figure 7 GRU Model.

There are many NLP techniques that automatically analyze human language and make it possible to represent it. The BERT algorithm, which consists of the initials of "Bidirectional Encoder Representations from Transformers", uses machine learning and artificial intelligence together.

The BERT given its architecture in Figure 8 is the most preferred natural language processing technique recently.

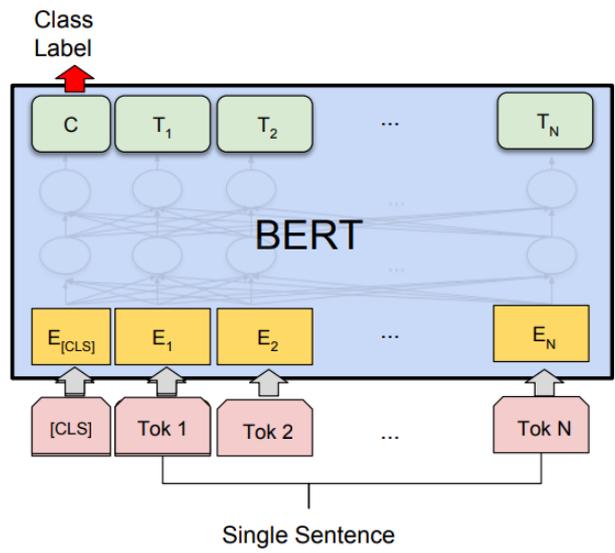


Figure 8 BERT Encoder

In order to analyze the data using the BERT algorithm, Python codes using Anaconda's Jupyter platform and open source deep learning libraries such as Keras neural network library and Tensorflow were used. Before analyzing the data, training and test data were created for the training of the machine, and this training data was tested. The results of the tests are shown in Figure 9.

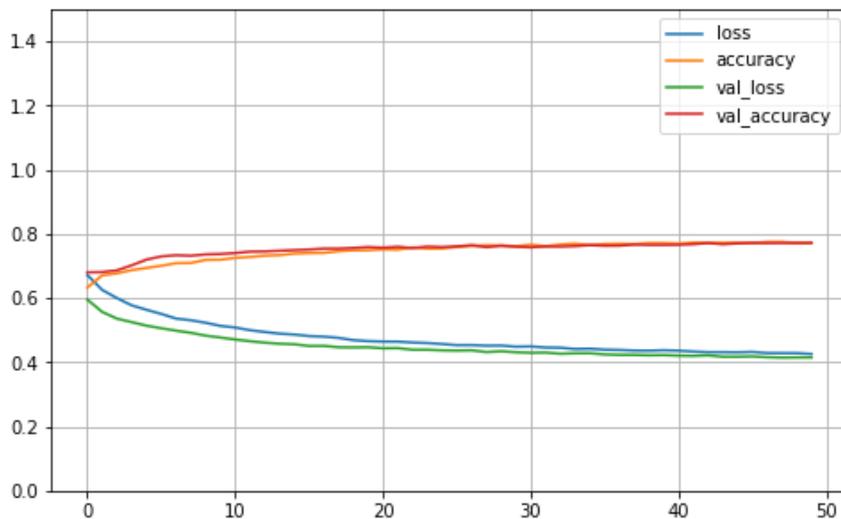


Figure 9. Test Results

Table 1 shows the results of three different architectures according to the performance metrics whose equations are given below.

$$Precision = \frac{TP}{TP+FP} \tag{1}$$

$$Recall = \frac{TP}{TP+FN} \tag{2}$$

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (3)$$

$$F_1 \text{ Score} = 2 * \frac{Precision*Recall}{Precision+Recall} \quad (4)$$

Table 1 Performance Values of Three Different Architectures

Performance Metrics	GRU	LSTM	BERT
Accuracy	0.78	0.77	0.78
Precision	0.79	0.78	0.77
Recall	0.93	0.92	0.76
F1 Score	0.85	0.84	0.77
AUC	0.88	0.88	0.88

While creating the model, the embedding layer was added before the GRU layers and word vectors were used as input. The resources in the LSTM method were used to train the model and generate the word vectors. In this study, sentiment score is the ratio of positive comments to total comments. Since short comments are mostly classified as neutral, the number of neutral comments is high. Short reviews are labeled neutral, as in the review phase, as they are not evaluated at all or are labeled neutral. Also, in LSTM/GRU experiments, the result of the prediction is produced between 0 and 1, thanks to a sigmoid. It is understood that the result is positive as these values approach 1, and negative as they approach 0. As can be seen in Figures 10 and 11, the rate of negative comments exceeds positive on some days.

The daily average of sentiments in Figure 10 and Figure 11 and the number of daily cases and deaths were compared. In comparison, it is observed that the sentiment value increases, that is, more positive comments are made on the days when the number of cases decreases. However, when compared with the number of deaths, the relationship between the emotional value and the daily number of deaths was not determined. This is really an expected result because the death is a later consequence of positive cases.

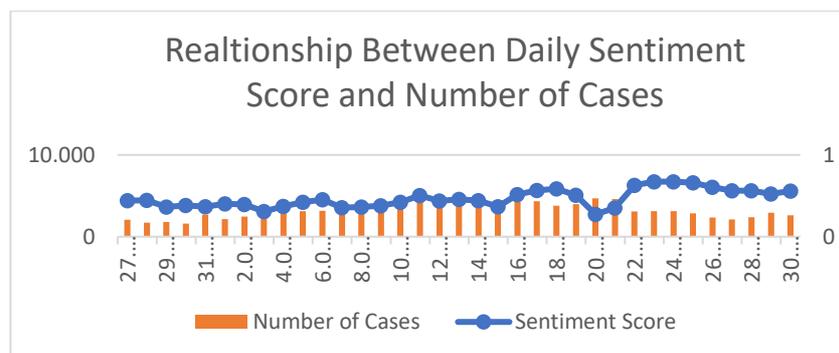


Figure 10 Relationship Between Daily Sentiment Score and Number of Cases

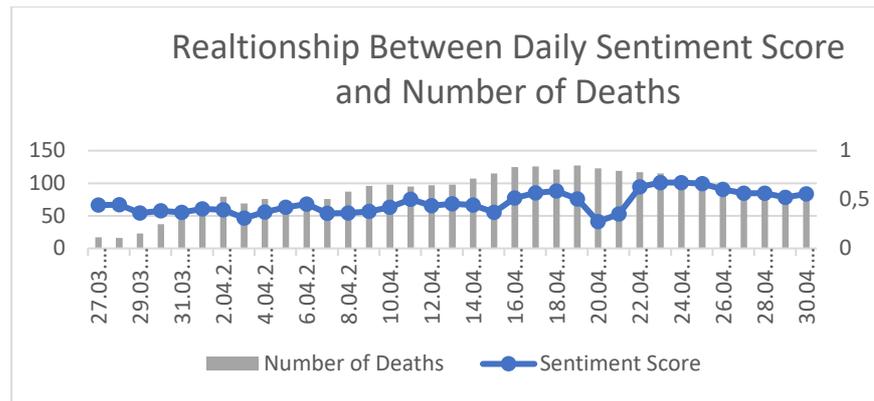


Figure 11 Relationship Between Daily Sentiment Score and Number of Deaths

5. Conclusions

Individuals who share their feelings and thoughts on social media channels have always been a great source of data for researchers observing society and individuals. In this paper, it is aimed to determine the feelings of individuals on Twitter about the number of positive cases and deaths related to the coronavirus pandemic, which started to affect the whole world at the beginning of 2020. As a dataset, only the comments made under the daily coronavirus tables announced by the Ministry of Health of the Republic of Turkey on its official Twitter account were used. The expressions in the obtained data set were subjected to various text preprocessing techniques and the emotional value of each tweet was calculated. Then, it has been tried to determine whether the values calculated using LSTM, GRU and BERT algorithms are related to the daily number of cases and deaths. The experimental results showed that there is a relationship between the number of positive cases per day and the sentiment score. On the other hand, since death is a later result of the positive case, it does not have an instant relationship with the sentiment score.

References

- [1] S. Kemp, "Digital 2020: 3.8 Billion People Use Social Media.", <https://wearesocial.com/blog/2020/01/digital-2020-3-8-billion-people-use-social-media>, 2020.
- [2] Ipsos, "Covid-19 Dönemi ve Evdeki Keşifler", https://www.ipsos.com/sites/default/files/ipsossia_trends_6nisan2020.pdf, Accessed August 20, 2020.
- [3] D. Murthy, *Twitter: Social Communication in the Twitter Age*. Cambridge, UK: Polity Press, 2018.
- [4] M. Szomszor, P. Kostkova and E. De Quincey, "# Swineflu: Twitter predicts swine flu outbreak in 2009", In *International Conference on Electronic Healthcare* (pp. 18-26). Springer, Berlin, Heidelberg, 2010.
- [5] J. Bian, U. Topaloglu and F. Yu, "Towards large-scale twitter mining for drug-related adverse events", In *Proceedings of the 2012 international workshop on Smart health and wellbeing* (pp. 25-32), 2012.
- [6] L. T. Nguyen, P. Wu, W. Chan, W. Peng and Y. Zhang, "Predicting collective sentiment dynamics from time-series social media", In *Proceedings of the first international workshop on issues of sentiment discovery and opinion mining* (pp. 1-8), 2012.
- [7] W.B. Claster, H. Dinh and M. Cooper, "Naïve Bayes and unsupervised artificial neural nets for Cancun tourism social media data analysis", In *2010 Second World Congress on Nature and Biologically Inspired Computing (NaBIC)* (pp. 158-163). IEEE, 2010.
- [8] B. Pang, L. Lee and S. Vaithyanathan, "Thumbs up? Sentiment classification using machine

- learning techniques", *arXiv preprint cs/0205070*, 2002.
- [9] R.M. Tong, "An operational system for detecting and tracking opinions in on-line discussion", In *Working Notes of the ACM SIGIR 2001 Workshop on Operational Text Classification* (Vol. 1, No. 6), 2001.
- [10] K. Ravi, V. Ravi. "A survey on opinion mining and sentiment analysis: tasks, approaches and applications", *Knowledge-based systems*, 89, 14-46, 2015.
- [11] M.L.B. Estrada, R.Z. Cabada, R.O. Bustillos, M. Graff. "Opinion mining and emotion recognition applied to learning environments", *Expert Systems with Applications*, 150, 113265, 2020.
- [12] İ. Sabuncu and M. Atmis, "Social Media Analytics for Brand Image Tracking: A Case Study Application for Turkish Airlines", *Yönetim Bilişim Sistemleri Dergisi*, 6(1), 26-41, 2020.
- [13] F. Ataman, Ö. Özgüner. "Determination of Social Media Users' Perceptions About "Black Friday" Activity via Sentiment Analysis", *Educational Research (IJM CER)*, 3(3), 361-371, 2021.
- [14] Z.H. Kilimci, "Financial sentiment analysis with Deep Ensemble Models (DEMs) for stock market prediction", *Journal of the Faculty of Engineering and Architecture of Gazi University*, 35(2), 635-650, 2020.
- [15] H.K. Küçükartal, "Twitter'daki Verilere Metin Madenciliği Yöntemlerinin Uygulanması", *Eskişehir Türk Dünyası Uygulama ve Araştırma Merkezi Bilişim Dergisi*, 1(2), 10-13, 2020.
- [16] M. Chandra, M. Reddy, S. Sehgal, S. Gupta, A.B. Buduru, P. Kumaraguru. "A Virus Has No Religion: Analyzing Islamophobia on Twitter During the COVID-19 Outbreak", In *Proceedings of the 32nd ACM Conference on Hypertext and Social Media* (pp. 67-77), 2021.
- [17] Z. Jianqiang, G. Xiaolin, Z. Xuejun. "Deep convolution neural networks for twitter sentiment analysis", *IEEE Access*, 6, 23253-23260, 2018.
- [18] A.A. Müngen, İ. Aygün and M. Kaya, "Finding the Relationship Between News and Social Media Users' Emotions in the COVID-19 Process", *Sakarya University Journal of Computer and Information Sciences*, 3(3), 250-263, 2020.
- [19] J. Samuel, G.G. Ali, M. Rahman, E. Esawi and Y. Samuel, "Covid-19 public sentiment insights and machine learning for tweets classification", *Information*, 11(6), 314, 2020.
- [20] F. Chen, Z. Yuan and Y. Huang, "Multi-source data fusion for aspect-level sentiment classification", *Knowledge-Based Systems*, 187, 104831, 2020.
- [21] Q. Lu, Z. Zhu, F. Xu, D. Zhang, W. Wu and Q. Guo, "Bi-GRU Sentiment Classification for Chinese Based on Grammar Rules and BERT", *International Journal of Computational Intelligence Systems*, 13(1), 538-548, 2020.
- [22] T. Wang, K. Lu, K.P. Chow and Q. Zhu, "COVID-19 Sensing: Negative sentiment analysis on social media in China via Bert Model", *IEEE Access*, 8, 138162-138169, 2020.
- [23] W. Li, H. Wu, N. Zhu, Y. Jiang, J. Tan and Y. Guo, "Prediction of dissolved oxygen in a fishery pond based on gated recurrent unit (GRU)", *Information Processing in Agriculture*, 8(1), 185-193, 2021.
- [24] B. Trstenjak, S. Mikac and D. Donko, "KNN with TF-IDF based framework for text categorization", *Procedia Engineering*, 69, 1356-1364, 2014.
- [25] Y. Bengio, I. Goodfellow and A. Courville, *Deep learning*, Massachusetts, USA: MIT Press, 2017.
- [26] M.E. Basiri, S. Nemati, M. Abdar, E. Cambria and U.R. Acharya, "ABCDM: An attention-based bidirectional CNN-RNN deep model for sentiment analysis", *Future Generation Computer Systems*, 115, 279-294, 2021.