Journal of Statistics & Management Systems
ISSN 0972-0510 (Print), ISSN 2169-0014 (Online)
Vol. 27 (2024), No. 3, pp. 605–633
DOI : 10.47974/JSMS-1028

Opinion mining and machine learning analysis : What emotions twitter data tell us about telemedicine?

Busra Saylan *
Songul Cinaroglu †
Department of Health Care Management
Faculty of Economics and Administrative Sciences (FEAS)
Hacettepe University
Beytepe 06800
Ankara
Turkey

Abstract

The Covid-19 pandemic has made telemedicine one of the most relevant topics in recent years, so data mining about telemedicine obtained from Twitter offers a unique opportunity. The motivation of this study is to identify the emotional groupings of Twitter data users' opinions for telemedicine using opinion mining techniques such as Sentiment Analysis combined with high-dimensional data classification and prediction methods. Data was collected from Twitter in August and September of 2021 related to telemedicine and official World Health Organization and World Bank documents for the years 2018 and 2021, respectively. Sentiment Analysis showed that 56.2% (n = 5351 tweets) of the sample had predominantly positive emotions toward telemedicine. Telemedicine, telehealth, and ivermectin are the most frequent words in the word cloud. Twitter data users' opinions consist of nine mood classes (SIL index = 0.80) and the differences between these classes are statistically significant in terms of positive ($p < 0.05$), negative ($p < 0.05$), and neutral ($p < 0.05$) emotions. Neural Network (AUC = 0.842, F1 = 0.492) and Random Forest (AUC = 0.841, F1 = 0.494) are the best predictors of Twitter users' mood classes compared with other machine learning techniques. The pythagorean tree generated by Random Forest showed that retweet is the best predictor of Twitter users' opinions and emotional classes towards

* *E-mail:* bsr.syln23@gmail.com (Corresponding Author)
ORCİD: https://orcid.org/0000-0002-1296-1824
† *E-mail:* songulcinaroglu@gmail.com, cinaroglus@hacettepe.edu.tr, cinaroglus@hotmail.com
ORCİD: https://orcid.org/0000-0001-5699-8402

telemedicine. Future studies will create big social media datasets for a deep understanding of the emotional classes of individuals towards telemedicine technologies.

Subject Classification: *62H30, 62M20, 68T10, 91C20.*

Keywords: *Telemedicine, Sentiment analysis, Opinion mining, Machine learning, Covid-19.*

1. Introduction

The concept of telemedicine, which emerged with the need to reduce people's face-to-face meetings, health service delivery, and access to these health services with the same access quality to benefit from health services, has come to the fore again as one of the most current problems experienced in recent years due to the Covid-19 pandemic [1]. Healthcare systems have faced a much greater burden due to the Covid-19 pandemic, which has raged around the world, demonstrating the urgent need for the application of technologies to facilitate improved connectivity between patients and healthcare providers [2]. Public health emergencies such as pandemics increase the demand for medical care and cause local capacities to be exceeded [3, 4]. For people to receive urgent health services during this period of the pandemic, telemedicine is the most appropriate service method, since it is capable of delivering health services beyond geographical location, time, culture, and social restrictions [5]. Telemedicine is defined by the World Health Organization (WHO) to improve the health of individuals and communities by using information and communication technologies to deliver health services where distance is a critical factor, to conduct research and evaluation by all health professionals, and to exchange valid information for the diagnosis, treatment, and prevention of diseases and injuries. Telemedicine, the delivery of healthcare services to patients at a distance, is not a new concept but was used in a number of proven health fields in Member States throughout the Western Pacific Region before the pandemic [1].

Essentially, telemedicine is remote medicine, which includes virtual reality, interactive imaging methods, and informatics applications. In the global e-health survey report published by the WHO (2016), telehealth is defined as health care provided when patients and health care providers are far from each other. It has been stated that it is aimed to inform patients about diagnosis and treatment by using information and communication technologies, to overcome geographical barriers, and thus contribute to the results of treatment. According to this report, 75% of the countries participating in the survey have tele-radiology practice, 50% have tele-health policy and 25% have state-supported tele-health programs [6].

The increase in the use of telemedicine in an emergency and pandemic is not a new phenomenon because it includes past outbreaks of coronavirus such as SARS-CoV (severe acute respiratory syndrome-associated coronavirus) and MERS-CoV (Middle East respiratory syndrome coronavirus) or PHEICs related to Ebola and Zika viruses. Recent evidence has shown that telemedicine has been used more heavily during local, national, and international crises as well as previous epidemics [7, 8]. The purpose of using telemedicine technologies during a global pandemic is not only to display, diagnose, and prioritize patients in their own relaxed environment, but also to reduce the need for physicians' prompt follow-up with patients, reducing hospital load and personal protective team use [4, 9]. Telemedicine has the capacity to help people who are not seriously ill get the supportive health care they need while minimizing their risk of exposure to other acute illnesses [10]. Applications of telemedicine are frequently used by health systems in the case of infectious diseases, such as pandemics. For example, telemedicine systems, which aim to meet the needs of people who are afraid of being exposed to the disease due to the pandemic, provide assurance and guidance for patients against the risk of contracting the disease, while reducing the patient density in emergency services, emergency care clinics, and primary care clinics. In addition, telemedicine can provide special care services for patients who want to access health services both at home and abroad [11]. However, the antecedents of telemedicine adoption expectations in different countries depend on income level and development level also based on factors such as patients' perception of service quality, institutional commitment, and satisfaction. It has been observed that there is a limited understanding of patient expectations and emotions regarding the identification of telemedicine needs and meeting these expectations in developing countries [12].

A study was conducted to determine the sentiment analysis of Twitter users for quarantine applications two weeks and four weeks after the quarantine was applied during the Covid-19 pandemic. According to this study, although there were great differences between the number of tweets containing positive and negative feelings towards quarantine practices on both dates, it was found that Twitter users generally exhibited a positive attitude towards quarantine practices [13]. According to a study conducted at New York University Langone Medical Center, it was observed that the number of emergency applications made through telemedicine services increased by 135% per day during the Covid-19 epidemic [14]. However, in a study in which sentiment analysis (SA) of tweets related to telemedicine was conducted during the Covid-19 pandemic, it was observed that users generally expressed emotions that were presented with a positive opinion against telemedicine during the pandemic period. This positive emotion sensitivity was followed by neutral and minor negative emotions. Thus, it has been found that the Covid-19 pandemic has increased the general interest in telehealth [15].

In a study examining the perceptions of healthcare professionals regarding telehealth platforms on Twitter before and after the Covid-19 pandemic period, it was observed that similar results were obtained in positive, negative, and neutral tweets in 2019 and 2020. As a result of the SA, it was determined that health professionals had a positive perception of telehealth in general, before and after Covid-19. In addition, according to other findings obtained from the study as a result of Epistemic Network Analysis (ENA), Covid-19 has shown that health professionals' perception of telescopic activity has changed from one of innovative application to one of safe access and maintenance, which correlates with increasing telescopic visits [16]. A survey was conducted with 711 adults from the USA to determine public opinion regarding the use of telemedicine for medicated abortion during the Covid-19 pandemic. According to the findings from this study, 44% of the adults surveyed supported the use of telemedicine for medicated abortion during the Covid-19 pandemic, while 35% opposed it. Most of the respondents who gave a positive opinion stated that medicated abortion was safe during the Covid-19 pandemic and that the practices of telemedicine were of comparable quality to face-to-face care [17].

International legal and regulatory agencies have also taken important steps to adapt to telemedicine technologies recently. Most European Union and Asian countries have resorted to legislation to integrate telemedicine systems into their own healthcare systems, providing more guidance on digital health technologies and cybersecurity prospects and expanding reimbursement options [18]. Such steps will be critical to increase access to healthcare for patients outside of physical healthcare facilities, in under-resourced areas, or who are overly afraid of accessing physical services for fear of exposure [11].

SA, also referred to as Opinion Mining (OM), was used to understand people's feelings, attitudes, thoughts, ideas, and evaluations towards entities such as products, services, organizations, problems, events, subjects, individuals, and their qualities [19]. The purpose of SA is to categorize the opinions in the text as “positive”, “negative” or “neutral” [20]. The emotion score is obtained as a result of the SA, and the positive score expresses satisfaction and happiness on behalf of the person, while the negative score expresses sadness, disappointment, and grief [21]. Today, where social media platforms are very important data sources, SA, and opinion mining techniques are widely used for large amounts of user-generated data. The machine learning (ML) approach is used effectively and reliably for the classification of these techniques [22]. A study was conducted to classify people’s suicide predictions and to determine their suicide prediction levels with tweets obtained from Twitter users. SA and classification techniques were used to better predict whether people are suicidal or not according to their tweets. It is seen that 86.42% of the tweets obtained from Twitter users were of negative polarity and for suicidal purposes. Sensitivity analysis by using Random Forest (RF) algorithm, to calculate the prediction of suicidal ideation, achieved an accuracy of 0.99% in estimating tweets containing suicidal ideation, compared to other classification methods [23]. Another study was conducted to estimate the probability of different age groups being affected by the coronavirus disease during the Covid-19 pandemic process and to use ML algorithms on the existing dataset. According to this study, it has been determined that the group that suffers most from Covid-19 is between the ages of 30-44, 45-59, and 60-74, the group with the highest death rate is between the ages of 60 and 74, and the group that is most likely to be contaminated is those over the age of 50. According to the results of the study, it was determined that the prediction models based on Random Forest Regressor and Random Forest Classifier, which are ML methods, outperform other prediction models [24]. According to another study, which aimed to classify non-textual tweets obtained from Twitter posts using sensitivity analysis using supervised learning techniques, it was determined that 36% of tweets without textual materials divided into classes and categories had positive moods. Considering the performances of the best estimators of the sensitivity classes obtained within the scope of the study, it was determined that the best method was The Support Vector Machine (SVM) with 68.15% [25].

In another study, in which the data obtained from Twitter was analyzed with SA to understand the emotions and behaviors of people who were under lockdown during the Covid-19 pandemic, and to identify people who developed suicidal and depression tendencies from the tweets within the scope of the study, 38.3% were neutral, 38% were positive and 23.7% were negative. According to the findings from this study, it was concluded that although Covid-19 has affected people's lives worse, the majority are still positive [26]. Lastly, a study was carried out in which it was aimed at developing an algorithm to reduce the negative consequences of fake news in social media and to automatically detect this news by using ML methods. In evaluating the performance of the algorithm developed within the scope of the study in detecting fake news, it was determined that SVM achieved the highest accuracy of 99.41% [27].

Telemedicine technologies are helpful during pandemic times to connect with health professionals and avoid getting infected. However, there is a lack of knowledge about individuals' emotions toward telemedicine. This study aims to determine the sentiment of Twitter data users' opinions on telemedicine during the Covid-19 pandemic. In order to evaluate the perceptions and attitudes of Twitter users regarding telemedicine during the pandemic, the positive, negative, or neutral moods of the data obtained from Twitter were examined by using emotion analysis techniques. Within the scope of this study, analyses will be supported by using visual graphics such as word clouds and heat maps. Also, as a next step, positive, negative, and neutral sentiments of Twitter data users' views towards telemedicine are determined by including a set of predictive variables through telemedicine country profiles, country income levels, retweets, and whether the tweet text contains the words Covid-19 and coronavirus. Performance evaluations will be made by using ML techniques such as Neural Network (NN), Random Forests (RF), Naive Bayes (NB), Support Vector Machine (SVM), and k-nearest neighbors (k-NN), to predict emotional groupings of Twitter data users' opinion. ML technique's performances are compared by using classification performance indicators. After determining the best-performing ML technique, the Pythagorean Tree was used to determine the best predictors of the emotional states of Twitter data users.

2. Methods and Materials

2.1 *Data collection*

Within the scope of the study, perceptions towards telemedicine during the Covid-19 pandemic were evaluated using social media posts on Twitter. A unique and original dataset containing a large amount of Twitter messages related to telemedicine was created. Using the RStudio library "rtweet", tweets with the keyword "telemedicine" were included in the study. Since the focus of the study is to collect only English tweets without any location restrictions, only tweets in English from Twitter through application programming interfaces (API) were included in the study.

Tweets about telemedicine were collected between 19 August 2021 and 30 September 2021. The dataset allows us to examine the effects of the increased need for telemedicine services during the Covid-19 pandemic and therefore to measure these perceived effects among Twitter users during the pandemic. While collecting data, it is aimed to reveal the effect of tweets by considering criteria such as the content of the tweet, the name of the user who tweeted, time, and location of the tweet. Data were collected from a total of 9516 tweets posted on topics related to "telemedicine" during the pandemic period.

In the second part of the analysis, newly created predictor variables are used to identify the best predictors of multi-class emotional groupings of Twitter data users' opinions towards telemedicine. Original Twitter text document, WHO, and WB resources are used to create this new dataset. Based on geographic locations of Twitter data users, the existence of country telehealth policy and strategy, country income levels, the existence of Covid-19 and coronavirus words in the Twitter text, and whether the tweets are retweeted are used as predictors of emotional classes.

2.2 *Data preprocessing*

During the data preprocessing steps in the Orange program, the methodological path used by Elghazaly et al. (2016) [28] was followed.

- Transformation

  In this step, the input data in the data set is transformed. Process steps such as the lowercase conversion of words in the input text, removal of accents in the text, parsing of html tags, and removal of url links in the text were applied.

- Tokenization

  During this process, the text is divided into words, sentences, etc. is converted. The method of separating into smaller components such as here, using the Regexp method, punctuation is skipped and the text is split according to the regular expression provided.

- Filtering

  In this process, firstly the words that are blocked from the text with Stopwords are removed. Then, the words in, on, etc. in the text have been removed. After that, a word selection is removed or retained in the next processing step. A regular expression method was applied to remove default punctuation.

- N-Grams

  During this step, n-grams were created from the markers, and the n-gram interval was determined as 1.

2.3 *Sentiment Analysis*

SA is one of the ML techniques that include the use of Natural Language Processing (NLP) [29]. SA aims to learn the attitudes, ideas, and feelings of individuals towards an event, entity, subject, or situation, it can also represent this entity, individual, event, or subject [30].

Although NLP and Linguistics have a deep-rooted research history, there were not many studies on the analysis of people's feelings and thoughts before the 2000s [19]. SA has become a very active research area since the 2000s [31]. SA answers many challenging research problems that have not been studied before, and for the first time in human history, now has a huge amount of convincing data on social media [19]. With the massive growth of social media in recent years, both individuals and institutions have begun to take these platforms, which are guided by the public in line with their own interests, more into consideration. SA, which is accepted as a sub-field of NLP, is also expressed as the process of understanding the opinions of the users on the subject of interest from the comments made by the users represented as unstructured texts on these social media platforms [32]. At this point, SA has come to the fore with the frequent use of blogs, news sites, review sites, and social media platforms such as Twitter for accessing large amounts of data sources. Because emotions are also closely related to the dissemination of information [31]. The high popularity of SA is parallel with the growth of social media usage. Therefore, SA is at the center of social media research. Moreover, SA does not affect NLP but also affects economics, political science, health, social sciences, and management sciences, which are affected by people's opinions and emotions [19].

The motivation of this study is to understand positive, negative, and neutral emotional groupings of Twitter data users' opinions and to identify the best predictors of positive, negative, and neutral Twitter mood classes towards telemedicine. Twitter API was used to collect tweets for this study. Orange ML and data mining program was used to process the text data. A SA of the tweets in the data set will be made and then, a positive, negative, or neutral sentiment score will be obtained for each word.

2.4 *Grouping of Twitter Data Users' Mood Classes in Terms of Telemedicine: A Heatmap Approach with K-Means Clustering*

In this study, the k-means clustering method was used for grouping of positive, negative, and neutral Twitter mood classes about telemedicine. The k-means algorithm is a non-hierarchical clustering method that divides objects into groups. The basic premise of the K-means clustering algorithm is to determine how many clusters to create first [33]. Data samples are assigned to the cluster with the closest center determined by the distance between these two points. Sample values from all files within the cluster are summed to calculate the center of each cluster. Thereafter, the process of dividing the data samples into respective clusters is repeated according to the updated cluster centers until the determined termination criteria are met. The method usually provides an initial configuration for more complex models [34].

Clustering a dataset is often difficult because the optimal number of clusters is uncertain. Cross-validation techniques such as k-fold cross-validation will assist in obtaining optimal grouping performance. In discrete Bayesian network analysis, k=10 is optimally used for k-fold cross-validation. With Bayesian network models, it is usually sufficient to use k=5 if the independent variables are independent of the variation of the variable values, especially when time and computational complexity are considered [35]. In our case, we incorporated k=5 fold cross-validation into the k-means clustering process while generating a heatmap and classified Twitter data users' opinions by means of positive, negative, and neutral moods towards telemedicine. To validate clustering results, Silhouette (SIL) index value was used as a cluster validity index score. The SIL index gives the average of pseudo-silhouettes across all clusters and is often used when comparing the results of different clustering techniques [36]. Additionally, Analysis of Variance (ANOVA) test was used to verify the robustness of the cluster validity index score. ANOVA test identified whether the difference between Twitter data user groups in terms of moods towards telemedicine was statistically significant.

In this study, we used the heatmap to visualize the clustering of positive, negative, and neutral Twitter mood classes toward telemedicine. In our case, the heatmap was created by using the k-means clustering technique, and k=5 fold cross validation incorporated, as mentioned before. A heatmap is a way to visualize values using a color gradient to visually represent values in individual cells, providing an overview of the largest and smallest values in both the matrix and within individual cells.

The heatmap matrix is typically clustered in such a way that it can be interpreted as a series of rows or columns rather than individual rows or columns [37].

2.5 *Machine Learning Predictors of Positive Negative and Neutral Twitter Mood Classes towards Telemedicine*

In this study, ML predictors such as NN, RF, NB, SVM, and k-NN were used to discover the best predictors of positive, negative, and neutral Twitter mood classes for telemedicine. For this stage of the analysis, a new data set consisting of four categorical independent variables was created.

During the first step of the data creation, two new categorical variables were generated based on the geographic locations of tweets. WHO eHealth Country Profiles Data (2018) [38] was used to create the first variable. Tweets were coded as 1 if they were from a country that implements telehealth policies throughout the country. The second categorical variable is based on the country income classification of the WB (2021) [39]. Tweets from low-income economies ($1045 or less) coded as 1, low-middle-income economies ($1045 to $4095) coded as 2, high-middle economies ($4096 to $12695) coded as 3, and high-income economies ($12696 and above) coded as 4. The number of retweets was taken as a basis for the creation of the third predictor variable. Depending on whether the tweets are retweeted or not, the variable has been made into two categories (0 = not retweeted, 1 = retweeted). Finally, since we aimed to analyze telemedicine emotional states during the pandemic process, we integrated Covid-19 into our study while creating the fourth variable. During the creation of this variable, we created a new variable with two categories (0 = no, 1 = yes) based on whether the words Covid-19 and coronavirus were mentioned in the tweet texts. The second part of the analysis compares the prediction performances of ML techniques and identifies the best predictors of the emotional groupings of Twitter data users' opinions using the newly created prediction variable set.

In this study, NN, RF, NB, SVM, and k-NN techniques were used and compared to predict the multi-class emotional groupings of the opinions of Twitter data users. NN is often used to solve classification and recommendation problems. This represents a mathematical and computational model based on biological NNs. This is an emulation of the biological nervous system [40,41]. The use of nonlinear statistical data modeling tools such as NNs can help model complex relationships among inputs and outputs or reveal hidden patterns in datasets [35]. In the input layer, the number of neurons represents the variables that are used to predict the output. In the output layer, the neurons represent the variables predicted [42]. The following formula is used to calculate the output of a neuron with n inputs (Equation 1) [43]. Y represents output, x_i represents input, w_i represents weight, b represents bias, and f represents a transfer function [42].

$$Y = f(\Sigma_{i=1}^{n} w_i x_i + b) \quad (1)$$

RF algorithms are ensemble learning algorithms that can be used to predict categorical or continuous response variables using classification and regression approaches. RF regression and classification models recursively divide the data into more homogeneous units called nodes, improving the predictability of the response variable for each tree by fitting a set of decision tree models to a set of data. Due to the split points being based on the values of the predictor variables, the variables used to split the data are important explanatory variables. Decision trees are fitted into a predefined number of preloaded datasets using RF. In categorical response, the predicted value is the mode of all available classes, while in continuous response, the predicted value is the average fit of all individual trees in each booted sample [44]. RF models offer excellent performance concerning a host of practical problems, in part because they are not prone to overfitting and do not suffer from noise in the dataset. Because it works fast, it usually exhibits a significant performance increase over many other tree-based algorithms [45]. RF has many advantages, but among the most important are: efficient implementation on large datasets, an easily saveable structure for future use of pre-built trees, and unmatched accuracy among existing algorithms [46]. Classification trees for binary outcomes are constructed using a training sample of "n" cases. This case has a set of covariates that are used to establish a tree-structured classification rule. Formally, the tree is grown by using this formula. (Equation 2) In this model, the regions Rj and the coefficients $\boldsymbol{\beta j}$ are estimated from the data. Rj is typically discrete and $\boldsymbol{\beta j}$ represents the mean of the Y values in Rj.

$$Y = \Sigma_{j=1}^{r} \beta_j I(x \in R_j) + \varepsilon \quad (2)$$

NB Classification method offers both supervised learning and statistical classification methods. It is a probabilistic model, which allows us to capture the uncertainties in modeling by establishing probabilities. It is named after Thomas Bayes, who proposed Bayes' Theorem in determining probability. NB helps to solve diagnostic and estimation problems [47]. NB is one of the most popular data mining algorithms for classification. Considering that the properties of a class are independent of other properties, the NB classifier infers the probability that a new sample belongs to a class [48, 49]. This assumption is made because multivariate probabilities must be estimated from the training data. Due to many values of the various attributes being either unavailable in the training data or not available in sufficient numbers, direct estimation of each multivariate probability is not feasible. NB circumvents this problem with the assumption of conditional independence [50]. Bayesian classification helps gain an understanding of and an evaluation of many learning algorithms. The method determines the exact probabilities for the hypothesis and is also resistant to noise in the input data [47]. (See Equation 3)

$$P(C \mid X) = \frac{P(X \mid C).P(C)}{P(X)} \tag{3}$$

P(C | X) is the posterior probability,

P(X | C) is probability,

P(C) is the pre-class probability,

P(X) is the predictive preliminary probability.

An SVM is a supervised machine-learning algorithm that creates a hyperplane between two types of structures, the maximizing distance between them. For example, it predicts the occurrence of heart disease by plotting the disease prediction features related to the multidimensional hyperplane, or it is used to classify the classes most appropriately by creating an approximation between two data sets [51]. The kernel trick has made SVM popular lately because it can deal with nonlinear classification problems [52]. After the creation of the hyperplanes, the SVM discovers the boundaries between the input classes and the input elements (support vectors) that define the boundaries, dividing a maximum margin hyperplane positive or negative training sample from a set of training samples labeled positively or negatively. As a result, the distance between the margin and the hyperplane is maximized [53]. When the input space is large and the input space is very high, the SVM classifier is highly generalizing without adding any prior knowledge. This is one of the most efficient and effective classifiers. It is used to solve binary classification and multiclass problems [54]. Consider a set of N distinct samples (x_i, y_i)

with $x_i \in \Re^D$ and $y_i \in \Re^d$ [55]. An SVM equation is modeled using the following formula (Equation 4). In this formula, K(x, x_i) represents the kernel function, while a and b represent the parameters and threshold of the SVM, respectively.

$$\Sigma_i a_i K(x, x_i) + b, i \in [1, N] \tag{4}$$

k-NN, which is based on the calculation of nearest neighbors, is one of the useful and well-known methods for supervised learning and is questioned by the similarity principle of Biau and Devroye (2015) [56] to predict a target output that an object or sample should have. The k-NN algorithm produces the target output of the new objects by using the closest samples in a feature space or by using the closest few objects in a training set [57]. Unlabeled objects are classified by calculating the distance between them and the tagged object and determining their "k" nearest neighbors. The accuracy of the classification usually depends on the selected value of "k" and will be more accurate than using the nearest neighbor algorithm [45]. This non-parametric algorithm is one of the simplest of all ML algorithms as it can solve both classification and regression problems. Despite its simplicity, k-NN is generally known to offer good performance in a wide variety of problems. It is very important to accurately calculate the similarity between samples to ensure good performance of k-NN. Its accuracy can be greatly affected by noisy or irrelevant features [57]. Data mining, image processing, etc. This algorithm, which is used in many applications in the fields, is used to calculate the distance of the training data closest to the object to be classified, using the Euclidean Distance method [58]. It is formulated as follows (Equation 5). The number of dimensions or features is n, x_i is the feature i in the data test, and y_i is the feature i in the data train [59].

$$dist(x, y) = \sqrt{\Sigma_{i=1}^{n} (x_i - y_i)^2} \tag{5}$$

Below are the performance methods used to compare ML techniques for telemedicine to determine Twitter users' mood classes. Accuracy is one of the popular metrics in multiclass classification and is calculated directly from confusion matrices. When calculating, the sum of the True Positive and True Negative items in the numerator and the sum of all entries of the confusion matrix in the denominator are taken into account while calculating the accuracy. In simple terms, accuracy calculates the probability that the model prediction is correct [60]. (See Equation 6)

$$Accuracy = \frac{TP+T}{TP+FN+FP+T} \tag{6}$$

As a ratio, Precision is the relevant items selected divided by all items selected [61]. True Positive are items that are labeled as positive by the model, False Positives are items that are rated as positive but are actually negative [60]. (See Equation 7)

$$\text{Precision} = \frac{TP}{TP+FP} \tag{7}$$

A Recall is defined as the number of relevant items selected over the total number of available items [62]. The False Negative contains elements that have been classified by the model as negative, but they are actually positive [60]. (See Equation 8)

$$\text{Recall} = \frac{TP}{TP+FN} \tag{8}$$

F1-score can be interpreted as the weighted average between Precision and Recall, where F1-score reaches its best value at 1 and its worst value at 0 [60]. (See Equation 9)

$$F1 = \left(\frac{2}{precision^{-1} + recall^{-1}}\right) = 2.\left(\frac{precision.recall}{precision+recall}\right) \tag{9}$$

AUC is most commonly understood as an integration method under a ROC curve. That is, the AUC represents the probability that a random sample with a positive label will rank higher than a random sample with a negative label. In the AUC formula, suppose that D^+ and D^- represent model prediction sets for positive and negative samples respectively (e.g. if $\hat{p}^{(i)} \in D^+$, then $\hat{p}^{(i)}$ is some probability in [0, 1], and the true label $y^{(i)}$ for instance $x^{(i)}$ is positive). $\tilde{\mathrm{I}}$ is a modified indicator function that returns 1 if the argument is positive, 0 if it is negative, and 0.5 if it is zero [63]. (See Equation 10)

$$AUC = \frac{1}{n_+ n_-} \Sigma_{\hat{p}^{(i)} \in D^+} \Sigma_{\hat{p}^{(j)} \in D^-} \tilde{\mathrm{I}}(\hat{p}^{(i)} - \hat{p}^{(j)}) \tag{10}$$

In this study, the Pythagorean Tree was created by using the RF algorithm to find the variable that best predicts the multi-class emotional groups of the views of Twitter data users. The Pythagorean tree, named after Pythagoras, is a plane fractal made of squares. In the Pythagorean tree, each triangle surrounds a right triangle, a symbol traditionally associated with the Pythagorean theorem. Its branches are generally straight and resemble veining [64, 65]. There are several visualizations of

Pythagorean forests based on a decision tree model called "Random Forest", where each visualization relates to a randomly constructed tree. The best type of tree structure is the one that has the darkest branches. This shows that few variables affect the branching of the tree. This technique applies to classification and regression problems [66]. The following section presents the results of this study.

3. Results

3.1 *Sentiment Analysis and Word Cloud*

In this study, SA was conducted for tweets collected from Twitter in August and September of 2021 related to telemedicine. The VADER (Valence Aware Dictionary for Sentiment Reasoning) model, a text SA method that takes into account the polarity (positive/negative) and intensity (strength) of the emotion, was used to analyze emotion. The weekly sensitivities of tweets about telemedicine without any location restrictions were monitored. After selecting only English tweets and applying the necessary data pre-processing steps on these selected tweets, 9516 tweets were selected for our study. Study results showed that 56.2% of the sample (n= 5351 tweets) had predominantly positive emotions, as seen in Table 1. Tweets containing negative emotions were recorded as 17,1 % (1620 tweets). As a result of our analysis, it was determined that 83% of the tweets recorded during the pandemic period were posted with positive or neutral emotions and Twitter data users' opinions towards telemedicine are mostly positive.

In addition, when the weekly word clouds of the collected tweets were examined, it was observed that all six weeks had similar word clouds. As seen in Figure 1, the 50 most frequently repeated words in the six-week data set were filtered out. However, when the word clouds of the six weeks were examined carefully, remarkable differences were observed in the word cloud for the dates corresponding to the 2nd week (August 26-September 1). As seen in Figure 1, when the 50 most frequently repeated words in the data set of the 2nd week were filtered out, it was seen that the word "ivermectin" stood out as one of the most frequently repeated words. The other most frequently repeated words are "telemedicine" and "telehealth", respectively, as in the word cloud of the five weeks. As it is known, ivermectin is a drug used to treat parasitic infestations. Note that, during the time, which is around week 2, the US Food and Drug Administration (FDA) has advised veterinarians and retailers to share

Word cloud for first 50 words in
week 1

Word cloud for first 50 words in
week 2

Word cloud for first 50 words in
week 3

Word cloud for first 50 words in
week 4

Word cloud for first 50 words in
week 5

Word cloud for first 50 words in
week 6

Figure 1

Word clouds for the first 50 words among tweets containing telemedicine in six weeks

with consumers important safety information about the dangerous misuse of animal ivermectin to prevent or treat Covid-19 in humans. Also, during this time, poison control centers across the US have reported a sharp increase in the number of people experiencing adverse health effects after taking animal ivermectin [67]. To the best of our existing knowledge, the FDA's explanations can be a reason for the high frequency of the word ivermectin.

Table 1

Sentiment analysis

Month	Total Tweets	Positive	%	Neutral	%	Negative	%
August (1st week)	2115	1330	24.9	485	19	300	18.5
August (2nd week)	1577	772	14.5	477	18.8	328	20.2
September (3rd week)	1662	846	15.8	442	17.3	374	23.1
September (4th week)	1416	759	14.1	450	17.7	207	12.8
September 5th week)	1585	912	17.1	392	15.4	281	17.3
September (6th week)	1161	732	13.6	299	11.8	130	8.1
Total Tweets	9516	5351	56.2	2545	26.7	1620	17.1

3.2 *Grouping of Twitter Data Users' Mood Classes in Terms of Telemedicine: A Heatmap Approach with K-Means Clustering*

Tweets about positive, negative, and neutral moods were classified using the k-means clustering method. Before clustering positive, negative, and neutral emotional states, study variables are standardized to avoid measurement unit differences. k=5 fold cross-validation was used for consistent selection of the number of clusters.

The optimal number of clusters was determined as nine with the SIL index value of 0.80, and the number of Twitter data users is balanced within each group (see Table 2). Nextly, the ANOVA test was used to detect if the nine clusters we obtained were different from each other, in terms of three emotional states, separately. ANOVA test results showed that the differences between nine Twitter data user groups' opinions in terms of positive ($F=12313,55$; $p<0.05$), negative ($F=11217,04$; $p<0.05$), and neutral ($F=10400,82$; $p<0.05$) mood classes are statistically significant.

The heat map corresponds to the relationships of positive, negative, and neutral moods related to telemedicine with the obtained clusters as in Figure 2. The heat map in the figure is colored according to the results of

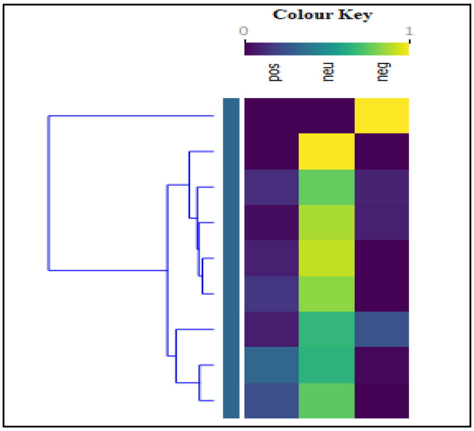

Figure 2

Heatmap analysis based on k-means clustering of negative, neutral, and positive moods related to telemedicine, showing hierarchical clustering on the left side

positive, negative, and neutral moods. Positive and negative moods are shown in blue and yellow, respectively, while neutral moods are shown in green. The rows in the heat map represent clusters, generated by k-means clustering and the columns represent moods. Note that, stronger moods are expressed in dark colors, while weaker moods are expressed in light colors. Also, nine groups of Twitter data users are presented in the dendrogram obtained from hierarchical clustering is visualized on the left side of the heatmap.

3.3 *Best Machine Learning Predictors of Positive, Negative, and Neutral Twitter Mood Classes towards Telemedicine*

In this analysis phase, a new dataset consisting of four independent categorical variables was used to determine the mood classes of positive, negative, and neutral groupings of Twitter users' opinions about telemedicine. Table 2 shows the basic statistics of the dependent and independent variable categories. It is seen that members of mood class 1 (C1) are higher (n=4265; 44.5%) compared with other classes. Moreover, the majority of Twitter data users are from high-income countries (83.2%), and 94.5 percent are from countries that have a national telehealth policy and strategy. Additionally, 90.7% of Twitter users did not use the words "Covid-19" and "coronavirus" in their tweets and 59.7% of them did not retweet.

Table 2

Descriptive statistics of variables to predict twitter data users' opinion mood classes

Dependent Variable									
Mood classes*	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**
n	4235	223	289	220	525	1261	2277	368	118
%	44.5	2.3	3	2.3	5.5	13.3	24	3.9	1.2
Independent Variables									
Country Income Classes**	**Categories**				**Country Telehealth policy*****	**Categories**			
	Low	**Low-middle**	**High-middle**	**High**		**Yes**		**No**	
n	37	1419	143	7917	n	8996		520	
%	0.4	14.9	1.5	83.2	%	94.5		5.5	
Covid-19 in the Twitter text****	**Yes**		**No**		**Retweet in the Twitter Text*******	**Yes**		**No**	
n	881		8635		n	3898		5618	
%	9.3		90.7		%	41		59	

* Data came from sentiment analysis results generated by emotional groupings of Twitter data users towards telemedicine.
** Data came from World Bank (2021) [39]
*** Data came from World Health Organization (2015) (Last updated: 2018) [38]
**** Data came from Twitter data analysis results and shows Covid-19 and coronavirus words in tweets.
***** Data came from Twitter data analysis results and shows the retweet status of tweets.

The purpose of creating this new dataset consisting of four independent categorical variables is to find out the best ML technique to predict clusters obtained from positive, negative, and neutral moods from Twitter users' opinions of telemedicine. ML techniques such as NN, RF, NB, SVM, and k-NN were used to make these performance comparisons. The performance values of the compared ML techniques are seen in Table 3.

Table 3 compares the prediction performance results obtained from NN, RF, NB, SVM, and k-NN algorithms. The performance of each model was evaluated in terms of Area Under the ROC Curve (AUC), Classification Accuracy (CA), F1, Precision, and Recall values on the dataset. Note that, AUC, CA, F1, precision, and recall values are preferable to compare the

Table 3

Comparison of machine learning techniques prediction models performances

Model	AUC	CA	F1	Precision	Recall
Neural Network	0.842	0.569	0.492	0.566	0.569
Random Forest (Number of trees=10)	0.841	0.573	0.494	0.577	0.573
Naive Bayes	0.838	0.564	0.524	0.599	0.564
Support Vector Machine	0.770	0.307	0.291	0.426	0.307
k-Nearest Neighbor	0.695	0.562	0.526	0.604	0.562
Abbreviations: AUC: Area Under the ROC Curve; CA: Classification Accuracy					

prediction performances of learning algorithms in the existence of multi-class predictor variables [68]. Study findings show that NN and RF perform slightly better than the other predictors, in terms of AUC and CA, the scores are more closer to 1.

According to these findings, NN had the highest prediction performance (AUC=0.842; CA=0.569), and RF (number of trees=10) is the second best predictor (AUC=0.841; CA=0.573). Whereas the NB (AUC=0.838; CA=0.564) is the third and SVM (AUC=0.770; CA=0.307) is the fourth, respectively. Among these ML techniques, k-NN (AUC=0.695; CA=0.562) showed the lowest prediction performance. The results verified that NN and RF have higher classification performances to predict multi-class Twitter data users' mood classes compared with other ML techniques.

As a final finding of this study, the Pythagorean tree was incorporated for data visualization and to see the most effective variable in determining the emotional classes of Twitter users' opinions. In Figure 3, the Pythagorean tree created by using the RF classification model (number of trees=10) obtained from the first cluster (C1), which represents the class of emotions that has a high number of members (n=4235), shows that retweet is the main decision node of the tree. Covid-19, income, and telehealth are the other variables that are effective in predicting the sentiment of Twitter users' views.

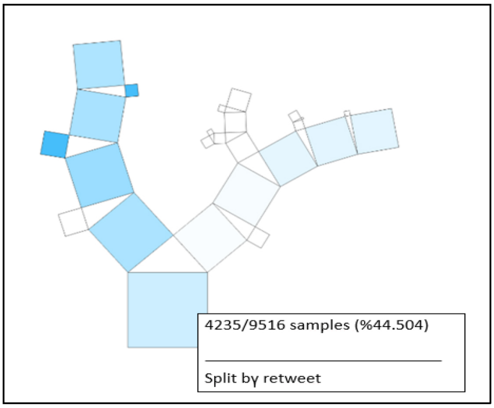

Figure 3

Pythagorean tree of emotion class C1.

4. Discussion

Understanding the public's thoughts and feelings towards health information technology is key to promoting communication technologies in healthcare. In times of pandemic, telemedicine is useful in alleviating healthcare access issues and creating closer links between rural and urban areas. The public's views on services provided in many fields, including health, and social media platforms, where these views are mostly shared, have recently become one of the most important sources of information in this context. Therefore, this study was conducted to examine the perceptions of Twitter users regarding telemedicine services during the Covid-19 pandemic. In this study, we observed the Twitter data of 9516 tweets related to telemedicine between August 19, 2021, and September 30, 2021, based on experimental analyzes and the results obtained. As a result of SA, we obtained similar results to the literature. According to the SA result, we found that Twitter users' perceptions of telemedicine during the pandemic were mostly positive. In other words, we can say that the need for telemedicine increased in parallel with the increasing health needs and quarantine conditions during the pandemic period.

After determining the positive, negative, and neutral emotions of Twitter users, an answer was sought to the question of which method was the most effective in predicting these mood groups. At this point, we used ML techniques to make performance comparisons to determine the answer to the question we were looking for. For this, we first created a new data set consisting of four independent categorical variables to determine mood classes and performed performance evaluations on this data set. When we examined the performance values, we found that NN and RF are the best learning algorithms to predict the emotional states of Twitter users towards telemedicine. Followingly, retweet is the best predictor of the emotional classes of Twitter users.

In this study, positive mood states for the telemedicine needs of the people in society during the Covid-19 pandemic were determined, and these emotional states were analyzed in more detail. This study is not without limitations. Within the scope of this study, only tweets in English were collected from Twitter, as it is the most frequently used language. Additionally, this study only verifies the emotional states of people who are using the Twitter platform. Further studies will search for other languages and other social media platforms. Finally, in this study, the Twitter dataset was created for six weeks period of time. Future studies will enhance the time period of the Twitter data generation process to create and analyze more comprehensive and big social media datasets.

5. Conclusion

Understanding public opinions and emotions towards health information technologies is the key to promoting communication technologies in healthcare. In times of pandemic, telemedicine is useful in alleviating healthcare accessibility problems and creating closer connections between rural and urban areas. This study collected data from Twitter which is a major social media platform during the pandemic period and explored Twitter users' emotions toward telemedicine by using opinion mining and ML techniques. Study findings showed that the dominant mood in tweets about telemedicine during the pandemic period was positive and Twitter data users consist of nine different groups in terms of positive, negative, and neutral mood classes. The further analysis incorporated ML techniques and detailed analysis of geographic regions and Twitter text documents obtained from Twitter users', to find the best predictors of Twitter users opinions mood classes. NN and RF are the best-performing learning techniques and retweet, Covid-19, and income are the best predictors of Twitter users' emotional classes towards telemedicine. Following current trends in telemedicine and exploring public opinions about health information technologies by using social media platforms is critical to inform healthcare policymakers during pandemics like Covid-19. In addition, it is recommended to work with a larger amount of data for which this study will be taken as a reference in future studies. Again, in future studies, to further expand the study, positive, negative, and neutral emotional states determined in detail are fear, joy, anxiety, sadness, panic, etc. emotions can be detected.

References

[1] World Health Organization. Implementing telemedicine services during COVID-19: guiding principles and considerations for a stepwise approach (2020). https://iris.wpro.who.int/handle/10665.1/14651. Accessed 25 September 2021.

[2] Pollack, C.C., Gilbert-Diamond, D., Alford-Teaster, J.A., & Onega T. Language and sentiment regarding telemedicine and Covid-19 on twitter: longitudinal infodemiology study. *Journal of Medical Internet Research*, 23(6), e28648: 1-10 (2021).

[3] Lurie, N., & Carr, B.G. The role of telehealth in the medical response to disasters. *JAMA Internal Medicine*, 178(6), 745-746 (2018).

[4] Hollander, J.E., & Carr, B.G. Virtually perfect? Telemedicine for COVID-19. *New England Journal of Medicine*, 382(18), 1679-1681 (2020).

[5] Annaswamy, T.M., Verduzco-Gutierrez, M., & Frieden, L. Telemedicine barriers and challenges for persons with disabilities: COVID-19 and beyond. *Disability and Health Journal*, 13(4), 100973 (2020).

[6] World Health Organization. Global diffusion of eHealth: making universal health coverage achievable: report of the third global survey on eHealth. Geneva,World Health Organization (2016).

[7] Ohannessian, R. Telemedicine: potential applications in epidemic situations. *European Research in Telemedicine/La Recherche Européenne en Télémédecine*, 4(3), 95-98 (2015).

[8] Smith, A.C., Thomas, E., Snoswell, C.L., Haydon, H., Mehrotra, A., Clemensen, J., & Caffery, L.J. Telehealth for global emergencies: Implications for coronavirus disease 2019 (COVID-19). *Journal of Telemedicine and Telecare*, 26(5), 309-313 (2020).

[9] Bokolo, A.J. Exploring the adoption of telemedicine and virtual software for care of outpatients during and after COVID-19 pandemic. *Irish Journal of Medical Science*, 190(1), 1-10 (2021).

[10] Portnoy, J., Waller, M., & Elliott, T. Telemedicine in the era of COVID-19. *The Journal of Allergy and Clinical İmmunology: In Practice*, 8(5), 1489-1491 (2020).

[11] Rockwell, K.L., & Gilroy, A.S. Incorporating telemedicine as part of COVID-19 outbreak response systems. *The American Journal of Managed Care*, 26(4), 147-148 (2020).

[12] Zobair, K.M., Sanzogni, L., & Sandhu, K. Expectations of telemedicine health service adoption in rural Bangladesh. *Social Science & Medicine*, 238, 112485: 1-10 (2019).

[13] Priyadarshini, I., Mohanty, P., Kumar, R., Sharma, R., Puri, V., & Singh, P.K. A study on the sentiments and psychology of twitter users during COVID-19 lockdown period. *Multimedia Tools and Applications*, 81(19), 27009-27031 (2021).

[14] Mann, D.M., Chen, J., Chunara, R., Testa, P.A., & Nov, O. COVID-19 transforms health care through telemedicine: evidence from the field. *Journal of the American Medical Informatics Association*, 27(7), 1132-1135 (2020).

[15] Champagne-Langabeer, T., Swank, M.W., Manas, S., Si, Y., & Roberts, K. Dramatic increases in telehealth-related tweets during the early COVID-19 pandemic: A sentiment analysis. *Healthcare,* 9(6), 634, 1-10 (2021).

[16] Ramanathan, P., & Jung, S. Healthcare professionals' perceptions of telehealth: analysis of tweets from pre-and during the COVID-19 pandemic. In Advances in Quantitative Ethnography: Second International Conference, ICQE 2020, Malibu, CA, USA, February 1-3, 2021, Proceedings, Vol. 1312, p. 390. Springer Nature (2021).

[17] LaRoche, K.J., Jozkowski, K.N., Crawford, & B.L., Haus, K.R. Attitudes of US adults toward using telemedicine to prescribe medication abortion during COVID-19: A mixed methods study. *Contraception*, 104, 104-110 (2021).

[18] Carrasqueiro, S., Ramalho, A., Esteves, A., Pereira, C., Martins, D., & Marques, L. Report on EU state of play on telemedicine services and uptake recommendations. Joint Action to Support the eHealth Network (2017).

[19] Liu, B. Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, 5(1), 1–167 (2012).

[20] Kharde, V., & Sonawane, P. Sentiment analysis of twitter data: a survey of techniques. *International Journal of Computer Applications*, 139(11), 5-15 (2016).

[21] Sv, P., Tandon, J., & Hinduja, H. Indian citizen's perspective about side effects of COVID-19 vaccine–A machine learning study. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 15,1-5 (2021).

[22] Ahmad, M., Aftab, S., Muhammad, S. S., & Ahmad, S. Machine learning techniques for sentiment analysis: A review. *International Journal of Multidisciplinary Sciences and Engineering*, *8*(3), 27 (2017).

[23] Rajesh Kumar, E., Rama Rao, K.V.S.N., Nayak, S.R., & Chandra, R. Suicidal ideation prediction in Twitter data using machine learning techniques. *Journal of Interdisciplinary Mathematics*, 23(1), 117–125 (2020).

[24] Patil, H., Sharma, S., & Raja, L. Study of impact of COVID-19 on different age groups using machine learning classifiers. *Journal of Interdisciplinary Mathematics*, 24(2), 479-487 (2021).

[25] Rawal, R., Bandil, D. K., & Nath, S. Data sparsity for twitter sentiment analysis in real-time from biased and noisy data. *Journal of Discrete Mathematical Sciences and Cryptography*, 24(8), 2403-2413 (2021).

[26] Sharma, S., & Sharma, S. Analyzing the depression and suicidal tendencies of people affected by COVID-19's lockdown using sentiment analysis on social networking websites. *Journal of Statistics and Management Systems*, *24*(1), 115-133 (2021).

[27] Malhotra, P., & Malik, S. K. Fake news detection using supervised machine learning techniques. *Journal of Information and Optimization Sciences*, 43(1), 7-15 (2022).

[28] Elghazaly, T., Mahmoud, A., & Hefny, H.A. Political sentiment analysis using twitter data. In Proceedings of the International Conference on Internet of Things and Cloud Computing, pp. 1-5 (2016).

[29] Li, J., & Hovy, E. Reflections on sentiment/opinion analysis. In: Cambria, E., Das, D., Bandyopadhyay, S., Feraco, A. (eds) A practical guide to sentiment analysis. Socio-Affective Computing, vol 5., pp. 41-59. Springer, Cham (2017).

[30] Venugopalan, M., & Gupta, D. : Exploring sentiment analysis on twitter data. In 2015 Eighth International Conference on Contemporary Computing (IC3), pp. 241-247. IEEE (2015).

[31] Ji, X., Chun, S.A., Wei, Z., & Geller, J. Twitter sentiment classification for measuring public health concerns. *Social Network Analysis and Mining,* 5(1), 13, 1-25 (2015).

[32] Sohrabi, M.K., & Hemmatian, F. An efficient preprocessing method for supervised sentiment analysis by converting sentences to numerical vectors: a twitter case study. *Multimedia Tools and Applications,* 78(17), 24863-24882 (2019).

[33] Bustamam, A., Tasman, H., Yuniarti, N., & Frisca Mursidah, I. Application of k-means clustering algorithm in grouping the DNA sequences of hepatitis B virus (HBV). In AIP Conference Proceedings, Vol. 1862, No. 1, p. 030134. AIP Publishing LLC (2017).

[34] Xie, H., Zhang, L., Lim, C.P., Yu, Y., Liu, C., Liu, H., & Walters, J. Improving K-means clustering with enhanced firefly algorithms. *Applied Soft Computing,* 84, 105763, 1-22 (2019).

[35] Marcot, B.G., & Hanea, A.M. What is an optimal value of k in k-fold cross-validation in discrete bayesian network analysis?. *Computational Statistics*, 36(3), 2009-2031 (2021).

[36] Starczewski, A. A new validity index for crisp clusters. *Pattern Analysis and Applications*, 20(3), 687-700 (2017).

[37] Metsalu, T., & Vilo, J. ClustVis: A web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. *Nucleic Acids Research*, 43(W1), W566-W570 (2015).

[38] World Health Organization. Atlas of eHealth country profiles 2015: The use of eHealth in support of universal health coverage. Based on the findings of the third global survey on eHealth. (Last updated: 2018) (2015).

[39] World Bank. Country and Lending Groups. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups. Accessed 01 October 2021 (2021).

[40] Jais, I.K.M., Ismail, A.R., & Nisa, S.Q. Adam optimization algorithm for wide and deep neural network. *Knowledge Engineering and Data Science,* 2(1), 41-46 (2019).

[41] Singh, Y., & Chauhan, A.S. Neural networks in data mining. *Journal of Theoretical & Applied Information Technology,* 5(1), 37-42 (2009).

[42] Amita, J., Singh, J.S., & Kumar, G.P. Prediction of bus travel time using artificial neural network. *International Journal for Traffic and Transport Engineering,* 5(4), 410-424 (2015).

[43] Beale, M.H., Hagan, M.T., & Demuth, H.B. Neural network toolbox user's guide. The MathWorks (2015).

[44] Everingham, Y., Sexton, J., Skocaj, D., & Inman-Bamber, G. Accurate prediction of sugarcane yield using a random forest algorithm. *Agronomy for Sustainable Development*, 36(2), 27, 1-9 (2016).

[45] Kodati, S., & Vivekanandam, R. Analysis of heart disease using in data mining tools orange and weka. *Global Journal of Computer Science and Technology C: Software & Data Engineering*, 18(1), 16-22 (2018).

[46] Kulkarni, A.D., & Lowe, B. Random Forest algorithm for land cover classification. *International Journal on Recent and Innovation Trends in Computing and Communication*, 4(3), 58-63 (2016).

[47] Parveen, H., & Pandey, S. Sentiment analysis on twitter data-set using naive bayes algorithm. In 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), pp. 416-419. IEEE (2016).

[48] Chen, S., Webb, G.I., Liu, L., & Ma, X. A novel selective naïve bayes algorithm. *Knowledge-Based Systems*, 192, 10536, 1-12 (2020).

[49] Kumar, R.S., Devaraj, A.F.S., Rajeswari, M., Julie, E.G., Robinson, Y.H., & Shanmuganathan, V. Exploration of sentiment analysis and legitimate artistry for opinion mining. *Multimedia Tools and Applications*, 81(9), 11989-12004 (2021).

[50] Zhang, L., Jiang, L., Li, C., & Kong, G. Two feature weighting approaches for naive Bayes text classifiers. *Knowledge-Based Systems* 100, 137-144 (2016).

[51] Iyer, A., Jeyalatha, S., & Sumbaly, R. Diagnosis of diabetes using classification mining techniques. *International Journal of Data Mining & Knowledge Management Process (IJDKP)*, 5(1), 1-14 (2015).

[52] Feng, W., Sun, J., Zhang, L., Cao, C., & Yang, Q. A support vector machine based naive bayes algorithm for spam filtering. In 2016 IEEE 35th International Performance Computing and Communications Conference (IPCCC), pp. 1-8. IEEE (2016).

[53] Nayak, J., Naik, B., & Behera, H. A comprehensive survey on support vector machine in data mining tasks: applications & challenges. *International Journal of Database Theory and Application*, 8(1), 169-186 (2015).

[54] Komal, K., & Sonia, D. : GLCM algorithm and SVM classification method for orange fruit quality assessment. *International Journal of Engineering Research & Technology (IJERT)*, 8(9), 697-703 (2019).

[55] Lo, S.L., Chiong, R., & Cornforth, D. Using support vector machine ensembles for target audience classification on Twitter. *PloS One*, 10(4), e0122855,1-20 (2015).

[56] Biau, G., & Devroye, L. Lectures on the nearest neighbor method, Vol. 246. Cham: Springer (2015).

[57] Triguero, I., García-Gil, D., Maillo, J., Luengo, J., García, S., & Herrera, F. Transforming big data into smart data: An insight on the use of the k-nearest neighbors algorithm to obtain quality data. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery,* 9(2), e1289, 1-24 (2019).

[58] Setiyaningrum, Y.D., Herdajanti, A.F., & Supriyanto, C. Classification of twitter contents using chi-square and k-nearest neighbour algorithm. In 2019 International Seminar on Application for Technology of Information and Communication (isemantic), pp. 1-4. IEEE (2019).

[59] Daeli, N.O.F., & Adiwijaya, A. Sentiment analysis on movie reviews using information gain and k-nearest neighbor. *Journal of Data Science and its Applications,* 3(1), 1-7 (2020).

[60] Grandini, M., Bagli, E., & Visani, G. Metrics for multi-class classification: an overview. ArXiv, 1-17 (2020).

[61] Fitri, V.A., Andreswari, R., Hasibuan, M.A. Sentiment analysis of social media Twitter with case of Anti-LGBT campaign in Indonesia using Naïve Bayes, decision tree, and random forest algorithm. *Procedia Computer Science*, 161, 765-772 (2019).

[62] Ling, J., Kencana, I.P.E.N., & Oka, T.B. Analisis sentimen menggunakan metode naïve bayes classifier dengan seleksi fitur chi square [Title in English: Sentiment Analysis Using Naïve Bayes Classifier Method with Feature Selection of Chi Square]. *E-jurnal Matematika* 3(3), 92-99 (2014).

[63] Kleiman, R., & Page, D. AUCμ: A performance metric for multi-class machine learning models. In International Conference on Machine Learning. Proceedings of Machine Learning Research (PMLR), pp. 3439-3447 (2019).

[64] Pourahmadazar, J., Ghobadi, C., & Nourinia, J. Novel modified pythagorean tree fractal monopole antennas for UWB applications. *IEEE Antennas and Wireless Propagation Letters,* 10, 484-487 (2011).

[65] Sathakathulla, A.A., & Akram, M. A note on cordial, edge cordial labeling of pythagoras tree fractal graphs. *International Journal of Scientific & Engineering Research*, 3(12), 1-5 (2012).

[66] Beck, F., Burch, M., Munz, T., Di Silvestro, L., & Weiskopf, D. Generalized Pythagoras trees for visualizing hierarchies. In 2014 International Conference on Information Visualization Theory and Applications (IVAPP), pp. 17-28. IEEE (2014).

[67] U.S. Food and Drug Administration (FDA). https://www.fda.gov/animal-veterinary/animal-health-safety-and-coronavirus-disease-2019-covid-19/cvm-letter-veterinarians-and-retailers-help-stop-misuse-animal-ivermectin-prevent-or-treat-covid-19. Accessed 30 August 2021 (2021).

[68] Deng, F., Huang, J., Yuan, X., Cheng, C., & Zhang, L. Performance and efficiency of machine learning algorithms for analyzing rectangular biomedical data. *Laboratory Investigation*, 101(4), 430-441 (2021).

Received July, 2022

Revised November, 2022