International Journal of Service Science, Management, Engineering, and Technology
Volume 13 • Issue 1

Multi-Class Sentiment Classification for Healthcare Tweets Using Supervised Learning Techniques

Brahami Menaouer, LABAB Laboratory, National Polytechnic School of Oran - M. Audin, Algeria*
https://orcid.org/0000-0003-0045-9797

Abdeldjouad Fatma Zahra, National Polytechnic School of Oran, Algeria

Sabri Mohammed, National Polytechnic School of Oran, Algeria

ABSTRACT

Social media has revolutionized the way people disclose their personal health concerns and express opinions on public health issues. In this paper, a new approach for multi-class sentiment classification using supervised learning techniques is presented. The aim of this multi-class sentiment classification is to assign the healthcare tweets automatically into predetermined categories on the basis of their linguistic characteristics, their contents, and some of the words that characterize each category from the others. Briefly, relevant health datasets are collected from Twitter using Twitter API; then, use of the methodology is illustrated and evaluated against one with only three different algorithms used to improve the accuracy of decision trees, SMO, and K-NN classifiers. Many experiments proved the validity and efficiency of the approach using datasets tweets, and it accomplished the data reduction process to achieve considerable size reduction with the preservation of significant dataset attributes.

KEYWORDS:

Medical Informatics, Sentiment Classification, Machine Learning, Health Tweet, Medical Decision Support System

1. INTRODUCTION

Today, Social Network Analysis (SNA) is commonly applied to investigating trends with studies of Social Media Analysis (SMA) and Data Mining (DM) used for such purposes (Khalifa et al. 2021; Wahi et al. 2014). In the last decade, Twitter has emerged as the most influential micro-blog service with twitter data source gaining considerable attention among researchers. Unlike many other social network services, twitter makes most user data world accessible. Additionally, twitter has numerous amounts of tweets, which mainly express opinions about a diversity of topics. These tweets may express valuable feedbacks and attitudes from patients about a specific disease or medical treatments. Twitter, unlike many other social networking services, makes most user data publicly accessible. These tweets, which express opinions about diverse topics, may also offer valuable feedback and reflect insightful attitudes from patients on specific diseases or medical treatments. O'leary (2015) observes that Twitter has become a critical social media tool with key capabilities, such as communication, building communities and collective action organizations. Yet, the fact that texts in social media are

DOI: 10.4018/IJSSMET.298669
*Corresponding Author

mostly written in colloquial language and both understanding and analyzing these texts is somewhat difficult in medical (epidemiological) context, further research attention in this area is needed, especially in text sentiment analysis. Specifically, sentiment analysis examines how sentiment is expressed in texts; briefly, sentiment systems may be used to identify sentiment categories from texts (Gridach et al. 2018; Tariyal et al. 2018).

In an earlier work, sentiment analysis is an automatic analysis technology of written or spoken speeches, aiming to extract subjective information like judgments, evaluations or emotions to detect the polarity of an opinion. According to (Kaur & Kautish, 2019; Bansal & Kaur, 2018; Rathi et al. 2018), sentiment analysis can be defined as the process of categorizing the opinions expressed through tweet to understand the user views about that topic. For (Yasen & Tedmori, 2019), sentiment analysis has been proposed as a component of other technologies. One idea is to improve information mining in document analysis by excluding the most subjective section of a text or to automatically propose internet ads for products that fit the viewer's opinion and removing the others. In general, sentiment analysis is performed utilizing one of two main approaches: the lexicon-based or unsupervised learning approach, in which rules extracted from the linguistic study of a language are applied to the sentiment analysis and the machine learning or supervised learning approach which relies on the famous machine learning algorithms to solve sentiment analysis as a classification task. Besides, the techniques for sentiment analysis on Twitter have been categorized into Lexicon Based, Machine Learning Based, Hybrid (Lexicon+Machine Learning) and Concept based (Ontology or Context) across pertinent literature (Kumar & Jaiswal, 2017).

In past decades, the use of information and communication technologies (ICTs) is included in all the activities namely healthcare and medicine development. This last is a knowledge-intensive domain obviously creates and uses large amounts of data and information. As well, we found that nearly 80% of searches through the search engine are related to health topics, symptoms, treatments. Currently, patients and citizens, in general, are increasingly using the Internet for searching health information and support. According to (Kamakshi, 2020), the use of online health communities is particularly popular among chronic patients. Surveys show that these patients significantly benefit from social interaction with peers and the sharing of knowledge and experiences (Kaoud, 2017). In the world that we are living today, sentiment analysis can also facilitate the healthcare industry to use reliable data for their growth by taking necessary measures. Sentiment analysis applies software to analyze the patient's tweets regarding their healthcare experiences regarding medicine, doctor, hospitals and others categories. It helps users as well as many healthcare organizations to understand their customers opinion and to take necessary measure to rectify the gaps. As well, information in social media platforms like Twitter, Facebook is also of great interest for researchers and professionals, as it allows for research the side effects of medicines, alternative treatments, pandemic spread (Covid-19) and quality and pollution of the environment. However, the quantity of information is so gigantic that it is difficult for the users to find the information that is really needed.

Furthermore, Machine Learning (ML) is the most effective method used in the field of data analytics in order to predict something by devising some models and algorithms (Brahami et al. 2022; Brahami et al. 2020; Anitha & Asha, 2018). In this field, there are many ML algorithms typically grouped by either learning style (i.e. supervised learning, unsupervised learning, semi-supervised learning) or by similarity in function (i.e. classification, regression, clustering, deep learning, etc.). To date, several researches have focused on using traditional classifiers, DM and ML based classifiers, such as Naive Bayes (NB), Decision Tree (DT), Sequential Minimal Optimization (SMO), K-Nearest Neighbor (K-NN), and Support Vector Machine (SVM) for Twitter classification (Madhuri, 2019; Asgarnezhad et al. 2018). Therefore, decision trees are recognized for their simplicity and effectiveness while the K-NN is easy to understand and robust to noisy training data. In addition, some advantages of SMO lies in the fact that solving for two Lagrange multipliers can be done analytically. In the literature, most of the available models classify user's tweets into positive and negative classes using traditional approaches. To improve the performance, it is needed to classify tweets in more than two classes, i.e., multi-class that are known as multi-class sentiment analysis. In the context of this study, we use these techniques for automatic multi-class sentiment on health tweets. Multi-class sentiment analysis aims to utilize this data by using more than two classes of sentiment. Multi-class sentiment analysis aims to automatically assign the health Tweets into predetermined categories based on their linguistic characteristics, contents, and some of the words that characterize the category from the others. For this, we used the healthcare-related tweets as an indicator about the quality of services provided by these health facilities to help to better understand the emotions expressed in the healthcare Tweets and to improve the provided services to patients. By training machine learning tools (Weka and CARTOCEL) with examples of emotions in tweets, machines automatically learn using Boolean modeling how to detect and classify sentiment according to categories (diseases, medical treatments, biological analysis, medications, epidemiology, and environment) without human input. The presented research in this paper addresses the following questions:

1. How to collect health-related comments from social media, such as Twitter?
2. What analysis is required to extract indicative knowledge from the collected tweets?
3. How extracted knowledge could be presented?

The remainder of this paper is organized as follows. Section 2 introduces the literature review on sentiment analysis methods and classification techniques proposed by various researchers. Section 3 presents the proposed methodology used for automatic sentiment classification for healthcare tweets. Section 4 presents the brief description of the classifiers (ID3, SMO, and K-NN). The experimental results and discussion are covered in Section 5. Finally, Section 6 concludes and summarizes the key findings and future works.

2. LITERATURE REVIEW

This section discusses the much work has been done in the area of sentiment analysis and text classification models. It is one of the most researched areas that often combine fields of Natural Language Processing (NLP) and Machine Learning (ML). In the literature, text classification is one of the important tasks of text mining. As well, it is a supervised ML problem. A lot of work has been done in the field of text classification. Here are various ways to gather the user's data and do the classifying textual data. A number of recent papers have addressed the classification of tweets most of them were tested against English and Arabic text (Bekkali & Lachkar, 2014). According to (Ur-Rahman & Harding, 2012), text classification is an important approach to handling textual data or information in the overall process of knowledge discovery from textual databases.

In earlier work, sentiment analysis is an automatic analysis technology of written or spoken speeches, aiming to extract subjective information like judgments, evaluations or emotions to detect the polarity of an opinion. According to (Bansal & Kaur, 2018), sentiment analysis can be defined as the categorizing process of the opinions expressed through tweets to understand the user's views about that topic. For (Yasen & Tedmori, 2019), sentiment analysis has been proposed as a component of other technologies. One idea is to improve data mining in text analysis by excluding the most subjective section of a text or to automatically propose internet ads for products that fit the viewer's opinion and removing the others. Besides, the techniques for sentiment analysis on Twitter have been categorized into Lexicon Based, ML Based, Hybrid (Lexicon+Machine Learning), and Concept-based (Ontology or context) across pertinent literature (Kumar & Jaiswal, 2017). In the work of (Velammal, 2019), sentiment classification is the process of analyzing the sentiment of a given text. It involves a combination of machine learning and natural language processing. For (Munikar et al. 2019), sentiment classification is a form of text classification in which a piece of text has to be classified into one of the predefined sentiment classes. Meanwhile, many Natural Language Processing (NLP) models have been proposed to solve the sentiment classification problem. According to (Bird et al. 2019), sentiment classification is an approach that can class the data into nominal labels or continuous polarities or score which map to their overall sentiment. Sankaranarayanan et al. (2009) have investigated the use of Twitter to build a news processing system, called TwitterStand, from Twitter tweets in order to capture tweets that correspond to late-breaking news. In the study done by (Foucault & Courtin, 2016) on the 103 French museums, defines the communication behavior of the 103 French museums that participated in 2014 in the Twitter operation: MuseumWeek using tweet classification. Gautam & Yadav, (2014) have proposed a set of techniques of ML with semantic analysis for classifying the sentence and product reviews based on twitter data using WordNet for better accuracy. In (Wang et al. 2016), authors have proposed a novel Hybrid Classification Algorithm (HCA) for descriptive sentiment analysis to understand students' problems and perks deeply. In this study, they have integrated both subjective analysis and data mining techniques to make the process descriptive. Similarly, sentiment classification is a process of dividing the target unit based on which the sentiments can be predicted. Another study by (Yelmen et al. 2018), have proposed a new method to improve the performance of classification of Turkish texts written in colloquial language, most particularly on social media using NLP methods. A study was conducted by (Gohil et al. 2018) in order to understand which tools would be available for sentiment analysis of Twitter health care research, by reviewing existing studies in this area and the methods they used. Similarly, (Srivastava et al. 2019) have presented a hybrid approach using Naïve Bayes and Random Forest on mining Twitter datasets. This hybrid methodology is illustrated and evaluated against one with only a Naïve Bayes classifier. Also, results show better accuracy and efficiency in the sentiment classification for the hybrid approach. A study was conducted by (Bird et al. 2019) in order to propose an approach to ensemble sentiment classification of a text to a score in the range of 1-5 of negative-positive scoring.

Over the years, many definitions have been given on sentiment analysis. In a previous study, Emadi & Rahgozar, (2019) have defined sentimental analysis as a process of computationally determining the opinion, or attitude of the writers as positive, negative, or neutral. In other words, (Swathi et al. 2019) said sentiment analysis is a sub-domain of opinion mining where the analysis is focused on the extraction of emotions and opinions of the people towards a particular topic from structured, semi-structured, or unstructured textual data. According to (Bonta et al. 2019), sentiment analysis is a type of data mining that deals with people's opinions through NLP, Computational Linguistics, and Text Analysis. There are mainly two approaches (Lexicon Based and ML) to extract the sentiment from given reviews and classify the result as positive or negative. To this end, many previous studies especially have used data sets (Tweets) and methods (ML and lexicon-based) for analyzing the opinions, sentiments. Otherwise, sentiment classification is a technique to classify the text expressed on social sites into various sentiment polarities (Rathi et al. 2018). In the meantime, many specific ML techniques are being used for sentiment classification and sentiment analysis. It is true that medical tweet data contains a huge volume of information in an unstructured format, for instance: sentiment-based N-gram machine learning, word embedding, and dependency-based rules (Madhuri, 2019). The work done by (Chakraborty et al. 2020), it is to bring out the fact that tweets containing all handles related to COVID-19 and WHO have been unsuccessful in guiding people around this pandemic outbreak appositely. The authors analyzed two types of tweets gathered during pandemic times. In the study of (Hassan et al. 2020), it is to (a) contribute to annotating Altmetrics data across five disciplines, (b) undertake sentiment analysis using Natural Language Processing and machine learning algorithms, (c) identify the best-performing model, and (d) provide a Python language for sentiment analysis of Altmetrics data. For this, (Tanulia & Girsan, 2019) have proposed to build a model by applying two methods, namely SVM and nonnegative matrix factorization in the process of predicting stock market movement using Twitter and historical data. To do that, the price and volume are taken from yahoo finance data, while topics and sentiment are taken from comments of each stock market in LQ45. In the work done by (Picassoa et al. 2019), it is to combines the technical and fundamental analysts' approaches to market trend forecasting through the use of ML techniques applied to time series prediction and sentiment analysis. Furthermore, the authors have aimed to develop a robust model able to predict the trends of a portfolio of stocks and to exploit its predictions in a trading strategy. In general, sentiment analysis is an active research area since it studied in many different levels namely information extraction and knowledge discovery from text using Natural Language Processing and DM techniques. In this respect, (Go et al. 2009) have proposed a novel approach for automatically classifying the sentiment of Twitter messages using distant super-vision processing. The authors classified the tweets using the feelings for example positive and negative tweets with positive feelings and negative feelings. Moreover, (Spencer & Uchyigit, 2012) have presented a tool for sentiment analysis of Twitter data named Sentimentor. This tool utilizes the Naive Bayes classifier to classify Tweets into positive, negative or objective sets. They have presented an experimental evaluation of dataset and classification results, findings are not contradictory with existing work. (Neethu & Rajasree, 2013) have used the two main techniques in sentiment analysis (symbolic techniques or knowledge base approach and ML techniques). The knowledge base approach requires a large database of redefined emotions and an efficient knowledge representation for identifying sentiments. In (Kasture & Bhilare, 2015), the authors presented the logical approach for extraction of the sentiment on widely used social networking sites. They analyzed the sentiments of the document using combinatory categorical grammar, lexicon acquisition and annotation, and semantic networks analysis. Likewise, (Foucault & Courtin, 2016) have presented a practical methodology for collect, pre-process, classify, summarize, and visualize the sentiment of the tweets. Then, they have evaluated a number of ML approaches and identified those most suitable to classifying public sentiment towards gun violence in light of the Sandy Hook school shooting. (Desai & Mehta, 2016) have proposed a novel Hybrid Classification Algorithm (HCA) for descriptive sentiment analysis to understand students' problems and perks deeply. For this, the authors have integrated both subjective analysis and DM techniques to make the process descriptive. (Kumar & Jaiswal, 2017) have conducted a study to contrast the two microblogging portals, Twitter and Tumblr determining the sentiment polarity using six supervised classification algorithms, namely, NB, SVM, Multilayer Perceptron, DT, KNN, and Fuzzy Logic. Moreover, (Daler et al. 2021) have presented the first-ever event detection approach for Urdu language text. Multiclass event classification is performed by popular deep learning models namely CNN, RNN, and DNN. The experimental work used a dataset that consists of more than 0.15 million (103965) labeled sentences. Madhuri, (2019) has proposed a framework for discovering sentiments from tweets of Indian Railways. This is a domain-specific framework that leverages business intelligence system through different classifiers such as SVM, ID3, Random Forest, and Naive Bayes. Besides that, (Al-Hadhrami et al. 2019) have used machine learning methods aimed at comparing the classification in sentiment analysis from the views of customers who have been written on Twitter. In addition, (Sirsat et al. 2019) have studied what highlights the usefulness of sentiment analysis along with the type of data that is being analyzed, the complex process involved in analyzing the data, the different approaches that can be used, and an experimental observation using the ML approach.

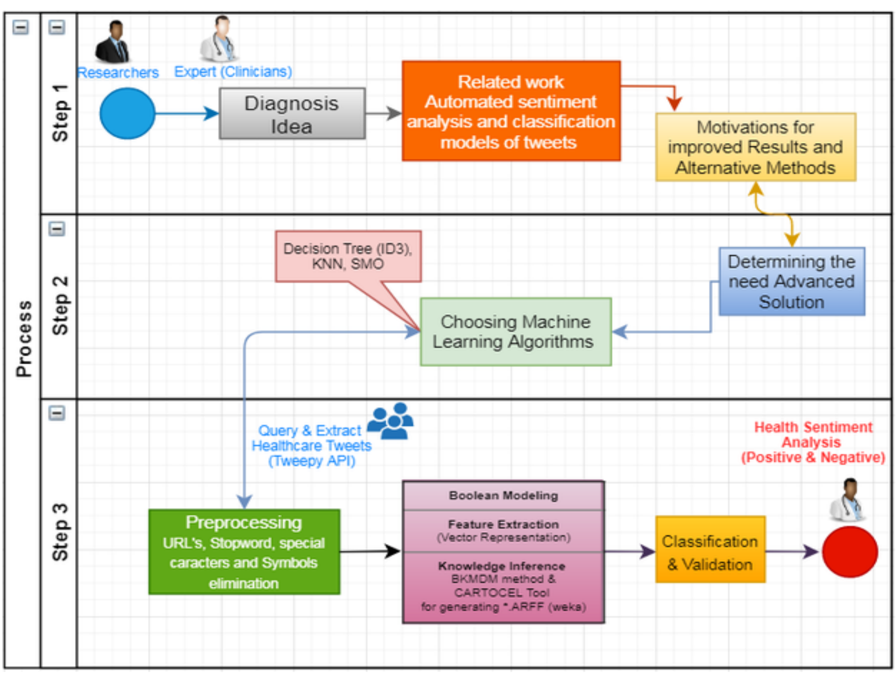
In the medical and health arenas and to our best knowledge, little work has been done to study sentiment analysis of social, and news media of medical and biomedical research using ML technologies. In this regard, some research efforts made by some researchers deserve mentioning. In fact, (Özçift, 2020) used supervised ML algorithms to predict the sentiment of the newly collected medical dataset in Turkish by developing an SVE algorithm to improve overall sentiment classification accuracy further. As well, (Divya, 2018) proposed an approach to analyzing user posts from health communities for knowledge discovery in order to help patients to find out the association among different drugs, diseases, and symptoms. Furthermore, this system will help doctors to find out the side-effects of different drugs so they can prescribe better drugs to other patients with similar diseases. In the study of (Sewalk et al. 2018), they provided a characterization of the patient knowledge sentiments across the United States on Twitter over a 4-year period. They have developed a set of 4 software components to automatically label and examine a database of tweets discussing the patient knowledge. The set includes a classifier to determine patient knowledge tweets, a geolocation inference engine for social data, a modified sentiment classifier, and an engine to determine if the tweet is from a metropolitan or nonmetropolitan area in the United States. (Zhao et al. 2015) proposed a novel bispace co-evolving framework named SimNest to integrate the complementary strengths of computational epidemiology and social media mining. Extensive experiments based on multiple states and flu seasons have demonstrated the advantages of integrating the respective strengths of computational epidemiology and social media mining. (Zaidan et al. 2015) presented a method to evaluate user interactivities in healthcare systems by a Facebook tool that measures statistical traffic on the Internet. They discussed the challenges and strengths of using such a platform as a tool for public health care systems from two different perspectives.

From the review of literature, sentiment analysis and classification are two highly challenging research areas that involve different tasks. The most studied tasks are sentiment classification and aspect sentiment classification. At the same time, due to the health-related experiences and medical histories these social media, namely tweets, provide for practitioners and patients, sentiment analysis tools had to be developed for the use in medical fields. Eventually, it is understood that ML techniques have the potential to mine sentiments associated with domain specific data sources. However, there is a need for evaluation of the performance of supervised learning methods for multi-class sentiment classification on tweets of healthcare and epidemiological so as to provide useful insights to make well-informed decisions.

3. PROPOSED METHODOLOGY

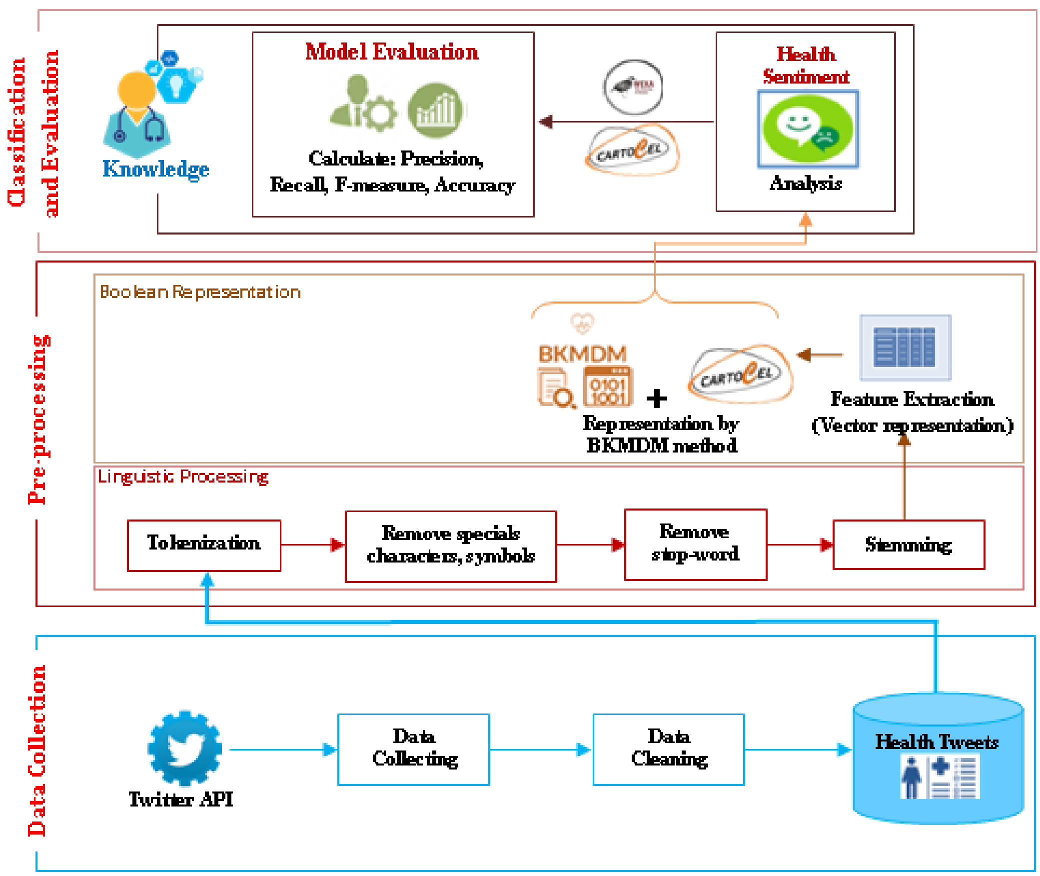
As per our objective and motivations, we implemented some background ideas and research efforts for automatic multi-class sentiment classification of healthcare tweets. As it can be observed in Figure 1 which shows the diagram of the main proposed contributions, the entire contributions are mainly divided into three steps: a collection of tweets that are potentially related to healthcare, data pre-processing, and classification. The following sections give out in detail a flow graph of the steps of our contribution (See Figure. 1).

Figure 1. Ideas and research efforts on the background of this study.



In the right context, the overall architecture of our proposed methodology is depicted in Figure 2 (See Figure. 2). The first step is the extraction of tweets. However, we developed a Python application using Twitter API that extracts a stream of tweets both during the desired time period and within a given region in Algeria. The second step is the preprocessing module which includes two models: linguistic processing that produces an index of words which is the vector representation and Boolean representation of the latter using the principle of BKMDM (Boolean Knowledge Management guided by Data Mining) method (Brahami et al. 2013; Brahami et al. 2020). The final step that is a recommender system is therefore proposed to achieve the aim of an effective classifying tweet that contains information about patients medical situation about taking certain diseases, and epidemics by using various machine learning algorithms (bagging, boosting, and stacking) to improve the accuracy of ID3, SMO, and K-NN classifiers. The main goal of these algorithms is to convert a weak learning algorithm into a strong learning algorithm, to reduce the false-positive rate.

Figure 2. Main Steps of the proposed health tweets classification methodology.



3.1. Collecting Health Tweets Dataset

The tweets are collected from the Twitter site using Twitter search Application Programming Interface (API). Secured access to this API is provided by a protocol OAuth. The API has the functionality to access real-time tweets in bulk and to investigate the content they hold (Singhvi & Srivastava, 2021;

Pang & Lee, 2004; Kaur et al. 2016). In addition, The Tweets contain three notations which are hashtags (#), retweets (RT), and account Id(@). As a case study, these latter were about different healthcare categories with several keywords for each category. As well, the data set consists of 600 tweets of different categories; each tweet was manually labeled based on their contents and the domain that it was found within. The tweets have been categorized into six categories mainly diseases, medical treatments, biological analysis, medications, epidemiology, and environment. Table 1 depicts the details of such a dataset (See Table. 1).

Table 1 Number of collected tweets per category.

Category name	Number of tweets
Diseases	100
Medical Treatments	100
Biological Analysis	100
Medications	100
Epidemiology	100
Environment	100
Total	**600**

Furthermore, the dataset has been divided into two parts: training and testing. The training data consist of 70% of the documents per category. The testing data, on the other hand, consist of 30% of the documents of each category. Finally, it took place over a period of seven days from the 22nd to the 30th of June, 2020.

3.2. Data Pre-preprocessing

In nearly all literature, pre-processing techniques are necessary, to acquire a more clean dataset. Basically, cleansing the dataset is actually a trial to improve text classification and increases the performance of the classification system by removing worthless data.

- Linguistic Processing: We present in detail the following operations which were performed as the data preparation tasks in order to extract features. For this, we parsed sentences into individual words (tokens) and we assigned an integer id for each possible token by using punctuation or white space as token separators. Then the tweet is checked for uniqueness, which ensures that the sentence does not contain any repeating words. Thereafter, the chosen health tweets are passed to the process of removing all irrelevant Twitter data. Furthermore, we removed URLs from tweets, because of URLs direct to extra information that was not a requirement for objectives in our study. The mention of other accounts with the @ sign were also removed including any other symbols or special character such as (&, #, /,%, *, etc..). In addition, a set of undesirable English stop words are excluded to improve and facilitate data processing. To remove stop words the proposed system has used the AFINN sentiment analysis lexicon and online English stop words list in order to remove the stop words. Moreover, identification of these stop words enables the decision-makers (e.g., Doctor and health technician) to retrieve information fast and makes the health tweets more powerful for information processing.

In one of the most recent studies, stemming is an essential process used in many fields of natural language processing like IR systems and textual classifiers. The objective of stemming is to arrive at a common stem by minimizing inflectional forms of words and derived related forms of a word. For Al-Shalabi et al. (2012), stemming is a process of linguistic normalization, in which the variant forms of a word are reduced. Stemming refers to a basic heuristic process that truncates the ends of words. It includes the removal of derivational affixes; e.g. the words '*vaccine*', and '*vaccinate*' are all turned into "*vaccine*". In addition, the words '*observe*', '*observes*', '*observer*', and '*observation*' of disease phenomena all could be stemmed to the word "*observe*". While stemming would shorten the words '*studies*' to '*studi*' and '*studying*' to "*study*", lemmatization would shorten both to "*study*". Under-stemming is when two words that should be stemmed to the same root are actually stemmed to different words.

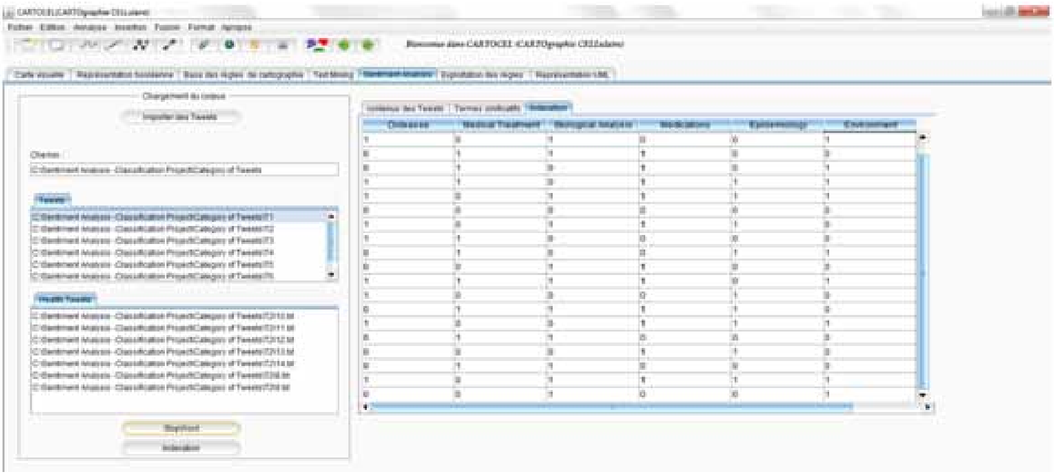
- Boolean Representation: Term Weighting is one of the pre-processing methods in order to help us display important words in a document collection for classification purposes (Alhanjouri, 2017). Most researchers claim that the aim of term weighting is to enhance text document representation as a feature vector or Vector Space Model (VSM). In VSM, each feature is represented as its term frequency or weight value using numbers. Therefore, term weighting is an important step that assigns a weight to indicate the importance of each related feature in the sample vector (Bozkurt et al. 2019). In contrast, transforming the preprocessing tweet into a feature vector is the first step in implementing the classifier (Abuelenin et al. 2017). In healthcare-related literature, selecting a good feature vector defines how strong a classifier can be. This is achieved using various approaches to term weighting, with the most well-known being the Boolean approach and TF*IDF (Term Frequency-Inverse Document Frequency) ranking technique which is the most commonly used weighting scheme in text classification (Chen et al. 2016). Vector representations of documents are then used in classification. For this reason, we use in this study the Boolean approach, which indicates the absence or presence of a word with Booleans 0 or 1 respectively. Table. 2 shows a sample of form files, using the Boolean representation (0,1) (see Table 2).

Table 2 Sample of format file includes the Boolean approach.

T1	T2	T3	T4	T5	....	Tn	Class
0	1	0	1	0	....	1	Diseases
0	1	1	0	1	....	0	Medical Treatments
1	0	1	0	1	....	1	Biological Analysis
0	1	0	1	0	....	1	Medications
0	0	1	1	1	....	0	Epidemiology
1	0	1	0	0	....	1	Environment

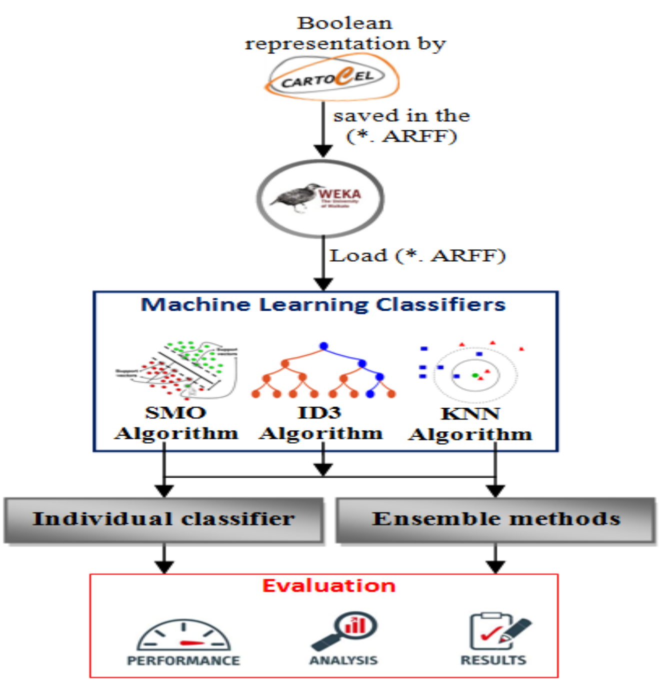
- Exploitation of the Vector Representation: After have been obtained a vector representation with binary weighting, the results are then saved in the (*.ARFF: (Attribute - Relation File Format) file format (extension), which is an extension file that can be read by WEKA (Waikato Environment for Knowledge Analysis) in order to exploit it with Data Mining software by using the BKMDM method through the CARTOCEL System (Brahami et al. 2013; Brahami et al. 2020). In general, these are open source knowledge mapping and data mining software mainly used for academic and research purposes. In general, these are open source knowledge mapping and data mining software mainly used for academic and research purposes. CARTOCEL proposes several machine learning algorithms and data preprocessing tools for knowledge learning from collect data, Boolean modeling, machine learning, and knowledge visualization. It provides the navigation and the interaction between the Boolean module and the graphics module so automatically. More recently, the result of the knowledge Boolean by BKMDM method is refined through induction rules obtained by this process of symbolic automatic learning to basic induction graph. Finally, the CARTOCEL system takes as its input the learning sample as a table of individuals/variables in order to supply a basis of classification rules during output, by applying the principle Boolean of the BKMDM method (See Figure. 3).

Figure 3. The Boolean representation of terms by CARTOCEL system.



The next step is to classify the tweets by Sequential Minimal Optimization (SMO), K-Nearest Neighbor (KNN), and Decision Tree (ID3 in Weka tool) classifiers using the cross-validation (CV) method. Eventually, we affirm that the final step is evaluating the accuracy results in two cases: individual classifier and combination algorithms case (See Figure. 4).

Figure 4. Outline for tweets Boolean representation and classification.



4. BRIEF DESCRIPTION OF THE CLASSIFIERS

Although ML has been applied in data analysis and related areas for decades. Solving problems in the medical domain using different tools, methods and techniques can be defined as Machine Learning on a known dataset. According to (Alsaeedi & Zubair Khan, 2019), classification methods have been developed, in the machine learning field, which uses different strategies to classify unlabeled data. In this paper, different machine learning classifier approaches have been adopted to achieve the highest possible performance on medical datasets and allows effective text classification. To predict tweets medical is considered as chronic diseases or pharmacovigilance, three predictive algorithms were chosen, namely: Decision Tree (ID3), SMO, and K-NN. In this field, we are using the most appropriate machine learning techniques namely DT, SMO, and KNN in order, on the one hand, to improve classification results in the domain of sentiment analysis and, on the other hand, to increase the efficiency and reliability of doctors' decisions (prediction). For the experiments, the Weka toolkit was chosen, as it provides an open-source scripting language specifically designed for classification, data analysis, and estimate the prediction accuracy. Finally, the results of each classifier have been compared, with the best classifier chosen for a more accurate tweets chronic disease prediction.

4.1. Decision Tree - ID3 Classifier

The decision tree is commonly used in operations research, specifically in decision analysis to help identify a strategy most likely to reach a goal, but is also a popular tool in machine learning (Sharma & Kaur, 2013; Ariestya et al. 2016). In addition, a decision tree is created by recursively selecting the best attributes to split the data and expanding the leaf nodes of the tree until a predefined stopping criterion is met (Wongsirichot et al. 2019). Primarily, a decision tree is like a flow chart that classifies instances depending upon the features. Each internal node represents the test case, branches show results of tests and leaf nodes hold the labels of classes (Kaur & Manik, 2018). Meanwhile, most algorithms developed for learning decision trees are derived from a basic algorithm that uses an up-down greedy search in the decision trees space. This method is indicated by the Dichotomiser 3 (ID3), C4.5, C5.0 and CART (Classification and Regression Trees) (Putra & Maulany, 2018).

In order to meet the goal of the work who is a construction of the decision tree, we used the ID3 algorithm method. ID3 (Inductive Decision Tree) is a heuristic tree that is used to construct a decision tree. Its principle consists of generating a succession of partitions by splitting nodes of the tree. Its objective is to optimize a criterion of information gain. From the sample of the learning method, ID3 is symbolic processing that begins the construction of the decision tree (Brahami et al. 2013; Fayyad et al. 1996):

- Choose the measurement uncertainty (Shannon or quadratic) ;
- Initialize the parameters of gain, info, and the initial partition S0;
- Apply the method ID3 to pass the partition S_t to S_{t+1} and to generate the decision tree;
- Finally, generate the prediction rules.

4.2. KNN Classifier

The k-nearest neighbor (K-NN) rule is one of the most common classification algorithms applied in many fields because it is very simple to understand and easy to design. Besides, it is mainly used when there is no prior knowledge about the distribution of data. Nevertheless, the execution time of k-Nearest Neighbors method increases dramatically when the size of the input dataset and the value of k is large (Bozkurt et al. 2019). The k-nearest neighbors' algorithm is a simple, supervised machine learning algorithm that can be used to solve both classification and regression problems. Specifically, the main aim of the KNN classification algorithm is to find the distance or similarity between the unclassified case and the training case with the known class label (Chen et al. 2016). The advantage of this algorithm is its simplicity in text categorization. It also works well with multi-class text classification. The main drawback of KNN is it necessitates with a large amount of time for categorizing entities, where huge data set, are inclined (Gridach et al. 2018; Al-Khasawneh, 2015).

As aforementioned, the similarity between the samples is determined by means of the Euclidean distance computation given as:

$$dist\left(p,q\right)=\sqrt{\sum_{k=1}^{k}\left(p_k-q_k\right)^2} \quad (1)$$

Where, p, q are the two samples and 'K' is the number of feature attributes.

Thus, the KNN classification can be done based on the nearest neighbor through the following representation:

$$q_x^{'}=\arg\max\sum_{p_k\in\propto\left(p_k\right)}f\left(q_x=L\right)\frac{1}{dist\left(p_x,p_k\right)} \ ;\ L\in\left\{L_1,L_2\right\} \quad (2)$$

In the equation (2) $dist\left(p_x,p_k\right)$ is the distance between the testing and training samples (Srivastava et al. 2019)

4.3. Sequential Minimal Optimization for SVM Classifier

Sequential Minimal Optimization (SMO) is a new algorithm for training Support Vector Machines (SVMs). It is iterative and finds use in solving optimization problems (Sanjudevi & Savitha, 2019). The latter assures that the amount of memory required for SMO is linear in the training set size, which allows SMO to handle very large training sets. For this reason, the main idea is derived from solving dual quadratic optimization problem by optimizing the minimal subset including two elements at each iteration. In the meantime, training a support vector machine requires the solution of a very large quadratic programming optimization problem (Nass et al. 2019; Jalali et al. 2017).

The result of the SMO algorithm is as follows (Singaravelan et al. 2015):

- Locate a Lagrange α_1 multiplier that infringes the Karush-Kuhn-Tucker (KKT) optimization conditions.
- Choose a second α_2 multiplier and optimize the pair (α_1, α_2).
- Repeat the above steps until merging.

When all the lagrange multipliers satisfy the KKT conditions (within a user-defined tolerance), the problem has been solved. Although this algorithm is guaranteed to converge, heuristics are used to choose the pair of multipliers so as to accelerate the rate of convergence (Zafar et al. 2019). This is critical for large data sets since there are $n(n-1)/2$ possible choices for α_i and α_j. Mainly, SMO's computation time is dominated by SVM evaluation, hence SMO is fastest for linear SVMs and sparse data sets (Pratibha et al. 2019).

5. EXPERIMENTAL RESULTS

After the accuracies of Sequential Minimal Optimization (SMO), K-Nearest Neighbor (K-NN), and ID3 (decision tree algorithm) have been got as individual classifier using cross-validation (10-fold) method, where the training data is divided randomly into (n) blocks, each block is held out once and the classifier is trained on the remaining (n-1) blocks, then three algorithms were used to improve these accuracies for health tweets classification those algorithms are Bagging, Boosting, and Stacking. As a shown in Table 3, the overall percentage accuracy for individual classifier: ID3, K-NN, and SMO (See Table. 3).

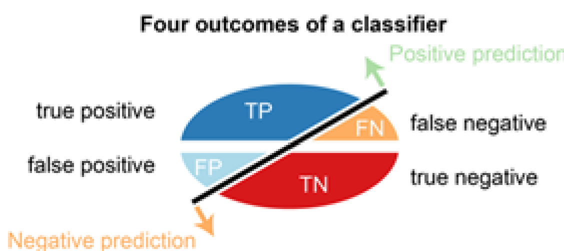
Table 3 Overall percentage accuracies for individual classifier.

Method	**10-fold cross validation**		
Individual classifier	***ID3***	***KNN***	***SMO***
Accuracy %	83.6	87.0	86.4

5.1. Evaluation of Results

Basically, the results obtained were analyzed by Weka and CARTOCEL tools based on several standard evaluation measures in order to identify the performance of the proposed approach in classifying the selected data. The measures are Accuracy (Acc), Precision (P), Recall (R), and F-measure (F). In short, accuracy as a measure is the number of samples that are correctly classified. According to (Othman & Al-Hamad, 2018), the precision calculation and recall are according to computing True Positive (TP: is the number of correct predictions of the positive instance or a number of medical entities that were identified correctly), True Negative (TN: is the a number of correct predictions of the negative instance), False Positive (FP: is the a number of wrong predictions of the positive instance or a number of medical entities that were detected by the system and were not present in the report) and False Negative (FN: is the number of wrong predictions of the negative instance or a number of medical entities that were present in the report but the system failed to detect them). Indeed, we present through Figure. 5 a different outcome of a two-class prediction and the rate of correctly predicted classes (See Figure. 5).

Figure 5. The rate of correctly predicted classes.



The detailed information of measure classification is shown in Table 4 (See Table. 4) and Figure 6 (See Figure. 6). Therefore, they display the F1-measure (*F*), and Recall (*R*) and Precision (*P*) for individual category (Diseases, Medical Treatments, Biological Analysis, Medications, Epidemiology and Environment) (See Table. 1 of Section 3.1) using cross-validation method.

In the right context, we can compute the precision as:

$$\text{Precision}\left(\text{P}\right) = \text{TP} \,/\, \text{TP} + \text{FP} \quad (3)$$

Also, we can compute the recall as:

$$\text{Recall}\left(\text{R}\right) = {}^{\text{TP}}\!/_{\text{TP} + \text{FN}} \quad (4)$$

Similarly, the F-measure combines precision and recall, the F-measure is used to calculate the performance of text classifiers as following equation:

$$\text{F} - \text{measure} = \frac{2\left(\text{Precision}*\text{Recall}\right)}{\text{Precision} + \text{Recall}} \quad (5)$$

And finally, the accuracy (overall success rate) is the number of correct classifications divided by the total number of classifications that we present by the following equation:

$$\text{Accuracy}\left(\text{acc}\right) = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (6)$$

In fact, it is the most common statistical measure that is used to evaluate the performance of a predictive model. It clarifies how well a classifier performs (Al-Khasawneh, 2015).

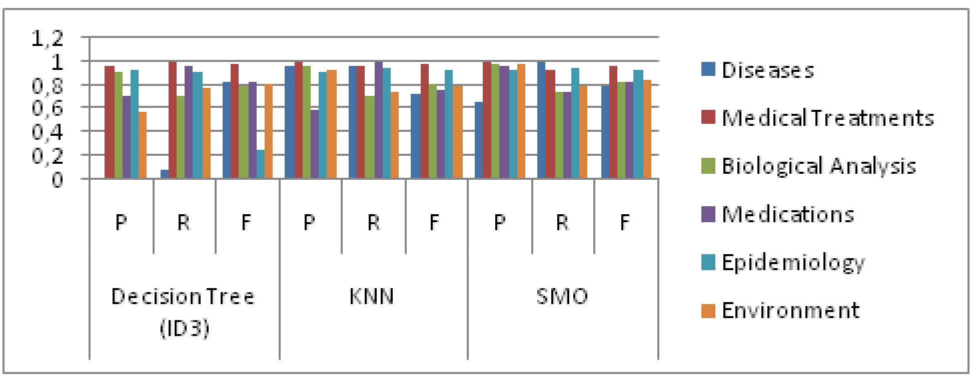
The error rate is also called the misclassification rate that shows incorrect classified cases, can be determined by:

$$\text{Error Rate} = \frac{\text{FP} + \text{FN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (7)$$

Table 4 overall accuracies P, R, and F for the individual categories

Classifier category name	ID3			KNN			SMO		
Cross-validation	***P***	***R***	***F***	***P***	***R***	***F***	***P***	***R***	***F***
Diseases	0.952	0.59	0.728	0.895	0.077	0.828	0.645	1	0.784
Medical Treatments	0.99	0.96	0.98	0.971	1	0.985	1	0.93	0.964
Biological Analysis	0.959	0.7	0.809	0.91	0.71	0.798	0.973	0.72	0.828
Medications	0.596	0.99	0.744	0.716	0.96	0.821	0.948	0.73	0.825
Epidemiology	0.913	0.94	0.926	0.919	0.91	0.915	0.931	0.94	0.935
Environment	0.579	0.78	0.811	0.93	0.74	0.801	0.980	0.79	0.845

Figure 6. P, R, and F for the individual categories.



5.2. The Cross-Validation Technique

In earlier works of machine learning, Cross-Validation is a method whose objective is to evaluate and compare learning algorithms. Likewise, cross-validation is the best way to stretch the validity of the manually annotated data since it enables it to be tested on a large number of documents. Furthermore, it consists of dividing the data into two segments: The first segment is used to learn or train a model and the second one is used to validate the model (Rahab et al. 2019). Usually, the advantage of this method is that all observations are used for both training and validation, and each observation is used for validation exactly once (Ravi et al. 2019; Mountassir et al. 2013). In the scope of this work (experiments), we have used 10-fold cross-validation. Using this technique the dataset is randomly split into 10 equal size subsets. Indeed, 10 cross-validation was used here since it is can be used to compare the performances of different predictive modeling techniques more objectively than simply using the in-sample error rates.

5.3. Improving Classification Accuracy

For improvement of the accuracy of health tweets multi-class classification, we used an ensemble of meta-algorithms that combine several machine learning algorithms in Weka into one predictive model in order to decrease variance (bagging), bias (boosting), or improve predictions (stacking).

- ***Improving Classification Accuracy using Bagging Algorithm:*** The first algorithm that we used to improve health tweets classification is bagging. Bagging (stands for Bootstrap Aggregating) is a way to decrease the variance of a prediction by generating additional data for training from an original dataset from where the predictions are combined by averaging or voting (Ridgeway, 2002). In general, this method is applied to decision tree algorithms, but it also can be used with other classification algorithms such as naïve bayes, nearest neighbor, etc. In this algorithm three classifiers ID3, KNN, and SMO have used it, to improve the accuracy of machine learning algorithms. Table 5 (See Table. 5) shows the accuracy for each classifier with the bagging algorithm. From Table. 5, we notice that the bagging algorithm achieves improvement in classification accuracy with ID3, KNN, and SMO classifiers, compared to the accuracy of individual classifier.

Table 5 The accuracies of ensemble methods (Bagging) for each classifier using 10-fold cross validation

Method	10-fold cross validation		
Ensemble	***Bagging***		
Classifier	***ID3***	***KNN***	***SMO***
Accuracy %	85.6	88.8	88.1

- *Improving Classification Accuracy using Boosting Algorithm:* The second algorithm that we used to improve the accuracy of health tweets classification is boosting. Boosting is a machine learning ensemble meta-algorithm for primarily reducing bias, and also variance in supervised learning, and a family of machine learning algorithms that convert weak learners to strong ones (Othman & Al-Hamad, 2018). Furthermore, boosting algorithm portrayed the idea of converting weak learning algorithms into an algorithm with high accuracy. In addition, the boosting is very little code and reduces variance, but it sensitive to noise and outliers. In this algorithm three classifiers ID3, KNN, and SMO are used with it, as shown in Table 6 (See Table. 6), we notice that the boosting algorithm achieves improvement in classification accuracy with ID3 and SMO classifiers, but with KNN classifier gave the same accuracy of individual classifier.

Table 6 The accuracies of ensemble methods (Boosting) for each classifier using 10-fold cross validation

Method	10-fold cross validation		
Ensemble	***Boosting***		
Classifier	***ID3***	***KNN***	***SMO***
Accuracy %	87.0	88.0	89.4

- Improving Classification Accuracy using Stacking Algorithm: The third algorithm that we used to improve the accuracy of health tweets classification is stacking. In the literature, stacking is an ensemble learning technique that combines multiple classifications or regression models via a meta-classifier or a meta-regressor. In summary, this stacking algorithm is very flexible and reduces both bias and variance. Therefore, the stacking algorithm takes more time than other classifiers as bagging and boosting to build a model (Ridgeway, 2002). Lastly, three basic models in this algorithm are used but the difference between them is the Meta classifier.

The first model in our work, ID3 classifier is used as a meta classifier; the second model KNN classifier is used as a meta classifier, and finally, SMO classifier is used as a meta classifier. In each of these models, we use a different number of base classifiers which we used in two cases. The first case is using two base classifiers ID3 and KNN. The second case is using three base classifiers ID3, KNN, and SMO. In this sense, we show through Table 7 (See Table. 7) the percentage accuracies for the stacking model with two base classifiers and a 10-fold cross-validation method.

Table 7 The percentage accuracies of stacking algorithm for each classifier in two base classifiers (ID3 and KNN)

Method	**10-fold cross validation**					
Algorithm	***Stacking***					
Meta classifier	***ID3***		***KNN***		***SMO***	
Base classifier	**ID3**	**KNN**	**ID3**	**KNN**	**ID3**	**KNN**
Accuracy %	86.4		87.8		88.4	

At the same time, we present through Table 8 (See Table. 8) the percentage accuracies for the stacking model with three base classifiers ID3, KNN, and SMO, and the 10-folds cross-validation method.

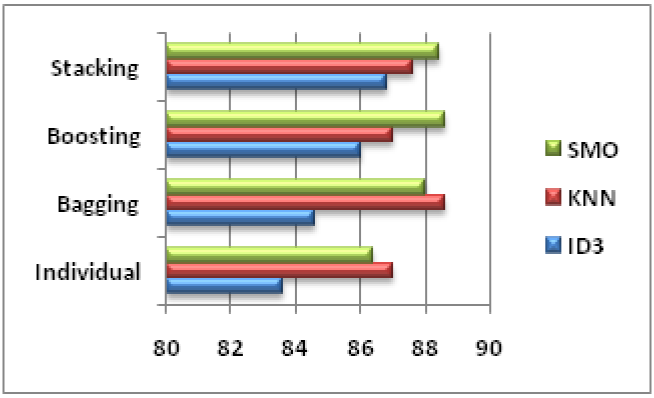
Table 8 The percentage accuracies of stacking algorithm for each classifier in three base classifiers (ID3, KNN and SMO)

Method	**10-fold cross validation**								
Algorithm	***Stacking***								
Meta classifier	***ID3***			***KNN***			***SMO***		
Base classifier	ID3	KNN	SMO	ID3	KNN	SMO	ID3	KNN	SMO
Accuracy %	86.9			87.7			88.6		

The result for the classification accuracy is significant and is given in Table 7 and 8. In fact, we notice that SMO classifier, when we use it as meta classifier achieving high accuracy in both two cases (two and three base classifiers) are 88.4% and 88.6% respectively, followed by KNN classifier with 87.8% and 87.7% respectively, and finally ID3 classifier with an accuracy of 86.4% and 86.9% respectively.

Eventually, the main motive of our work was to compare ensemble methods with each classifier and individual classifiers. In addition, we show in the current study increase of in accuracy for each classifier. In this sense and through Figure 7 (See Figure. 7), we noticed that the ID3 classifier; gradually increasing accuracy with each algorithm (Bagging, Boosting, and Stacking) respectively. For KNN classifier, the accuracy reached up 88.6% in the bagging algorithm, then the accuracy reached down to 87.0% in the boosting algorithm, and finally with stacking algorithm the accuracy reached 87.60%. Moreover, the SMO classifier achieves with boosting and stacking algorithms high accuracy reached up to 88.6% and 88.4% respectively, while with bagging with an accuracy of 88.0%. Finally, the ID3 classifier achieves high accuracy with boosting and Stacking algorithms reached up to 86.0% and 86.8% respectively.

Figure 7. A comparison between ensemble methods with each classifier and individual classifiers using 10-fold cross validation.



In summary, our claim holds true, that social media information has an important and large source of information about individuals. For this, sentiment analysis using social media data will thus provide valuable insights on attitudes, perceptions, and behaviors for decision-making for health professionals and patients. Besides that, automated machine learning-driven sentiment analysis could help health professionals, policymakers, and the government to understand and identify rapidly changing epidemiological risks in the population. This study showed that the results are better when applying many algorithms rather than applying a single algorithm on a social media healthcare dataset. CARTOCEL (Brahami & Matta, 2019) and WEKA (Dhakate et al. 2014) are chosen in this work. They are some of the other popular tools used for data analysis. As well, this work shows that these patients significantly benefit from social interaction with peers and the sharing of knowledge, experiences and support on the disease, medical treatments, medications, and environment. In contrast, social media has not been useful enough to help people during epidemic (infectious) diseases (for example SARS, MERS, Ebola, and currently COVID-19 outbreak). As well, the automatic classification according to six categories not only expresses sentiment for healthcare problems but also supports the opinion with strong reasons which make useful information to the public. In this sense, the proposed methodology and the findings can also be further extended to similar local and global disease insights generation in the future. Eventually, past research has also identified the use of Twitter data analytics for different diseases, namely infectious disease, indicating a mature stream of thought towards using social media data to help understand and manage infections, crisis and risk scenarios.

6. CONCLUSION

Twitter Sentiment Classification plays an essential role in the attitudes of users and emotional states. People tend to express their feelings freely, which makes Twitter an ideal source for accumulating a vast amount of opinions towards a wide spectrum of topics namely healthcare. Moreover, the complexity of natural languages and the difficulty of understanding how humans express their feelings is a challenging task in multi-class classiðcation. This amount of information offers huge potential and can be harnessed to receive the sentiment tendency towards the topic of healthcare tweets. The main contribution of this paper is to improve an automatic sentiment multi-class for health tweets dependent on machine learning techniques. In our case study, the health tweets are classified into one of some predetermined categories mainly (Diseases, Medical Treatments, Biological Analysis, Medications, Epidemiology, and Environment) with over 600 tweets collected. For this goal, we employed Decision Tree (ID3), KNN, and SMO classifiers in order to extract opinion targets. Furthermore, the ensemble methods or combination methods (bagging, boosting, and stacking) were used to improve the multi-class accuracy. Hence to obtain the best results, these methods require a large corpus to allow better learning. The proposed approach helps the help health professionals and policymakers to identify the major national issues providing a golden opportunity to take positive steps to resolve the problems of the masses.

A limitation of the study is the use of a limited number of our current tweets datasets. We intend to make our model more robust and accurate by using more such tweets from other social media platforms. Indeed, traditional classification models work better for small datasets. When data size increases, these models' accuracy goes down. In addition, we intend to test our model on other sentiment categories of the Human Being (patient) namely Disorder, Symptoms, Depressed and Fear, Worry, and Body Measurement. Besides that, we intend to test our model on English, French and Arabic tweets coming from unspecified locations.

As for future research, multi-class classification with a deep learning approach (CNN and RNN) is another great area to explore in particular for feature-based sentiment classification to determine the degree of sentiment and to analyze the user opinion. Likewise, sentiment classification with Fuzzy logic in order to determine the degree of sentiment which allows certain degrees of membership of values between 0 and 1. Currently, there are various classification methods, such as fuzzy K-Nearest Neighbor (F-KNN) and Fuzzy Decision Trees (FDT). Similarly, sentiment classification a larger Twitter Data (Data Lakes) is another great area to explore and we intend to use big data analytics. Furthermore, the current machine learning technique can be replaced with more powerful methods/ algorithms namely deep learning approaches and hybrid methods to compute, analyze and predict the results much faster and accurate with a minimum error rate.

FUNDING AGENCY

The publisher has waived the Open Access Processing fee for this article.

REFERENCES

Abuelenin, S., Elmougy, S., & Naguib, E. (2017). Twitter Sentiment Analysis for Arabic Tweets. *Proceedings of the International Conference on Advanced Intelligent Systems and Informatics*, 467–476.

Al-Hadhrami, S., Al-Fassam, N., & Benhidour, H. (2019). Sentiment Analysis Of English Tweets: A Comparative Study of Supervised and Unsupervised Approaches. *Proceeding of the 2nd International Conference on Computer Applications & Information Security (ICCAIS)*, 1–6. doi:10.1109/CAIS.2019.8769550

Al-Khasawneh, A. (2015). A Method for Classification Using Data Mining Technique for Diabetes: A Study of Health Care Information System. *International Journal of Healthcare Information Systems and Informatics*, *10*(3), 1–23. doi:10.4018/IJHISI.2015070101

Al-Shalabi, R., Ababneh, M., Kanaan, G., & Al-Nobani, A. (2012). Building an Effective rule based light Stemmer for Arabic Language to improve Search Effectiveness. *The International Arab Journal of Information Technology*, *9*(4), 312–316.

Alhanjouri, M. (2017). Pre Processing Techniques for Arabic Documents Clustering. *International Journal of Engineering and Management Research*, *7*(2), 70–79.

Alsaeedi, A., & Zubair Khan, M. (2019). A Study on Sentiment Analysis Techniques of Twitter Data. *International Journal of Advanced Computer Science and Applications*, *10*(2), 361–374. doi:10.14569/IJACSA.2019.0100248

Anitha, A. V., & Asha, A. (2018). Improving Prediction Accuracy Using Hybrid Machine Learning Algorithm on Medical Datasets. *International Journal of Scientific and Engineering Research*, *9*(10), 1461–1467.

Ariestya, W. W., Praptiningsih, Y. E., & Supriatin, W. (2016). Decision Tree Learning Untuk Penentuan Jalur Kelulusan Mahasiswa. *Jurnal Ilmiah FIFO*, *8*(1), 97–105. doi:10.22441/fifo.v8i1.1304

Asgarnezhad, R., Monadjemi, S. A., Soltanaghaei, M., & Bagheri, A. (2018). SFT: A Model For Sentiment Classification Using Supervised Methods On Twitter. *Journal of Theoretical and Applied Information Technology*, *96*(8), 2242–2251.

Bansal, P., & Kaur, R. (2018). Twitter Sentiment Analysis using Machine Learning and Optimization Techniques. *International Journal of Computers and Applications*, *179*(19), 5–9. doi:10.5120/ijca2018916321

Bekkali, M., & Lachkar, A. (2014). *Arabic Tweets Categorization based on Rough Set Theory*. Presented at *Third International Conference on Data Mining & Knowledge Management Process (CDKP-2014)*, Dubai, UAE. doi:10.5121/csit.2014.41109

Bird, J. J., Ekárt, A., Buckingham, C. D., & Faria, D. R. (2019). High Resolution Sentiment Analysis by Ensemble Classification. *Intelligent Computing - Proceedings of the Computing Conference*, 1–14.

Bonta, V., Kumaresh, N., & Janardhan, N. (2019). A Comprehensive Study on Lexicon Based Approaches for Sentiment Analysis. *Asian Journal of Computer Science and Technology*, *8*(S2), 1–6. doi:10.51983/ajcst-2019.8.S2.2037

Bozkurt, F., Çoban, Ö., Günay, F. B., & Yücel Altay, Ş. (2019). High Performance Twitter Sentiment Analysis Using CUDA Based Distance Kernel on GPUs. *The Journal Tehnički Vjesnik – Technical Gazette, 26*(5), 1218–1227.

Brahami, M., Abdeldjouad, F. Z., Sabri, M., Semaoune, K., & Matta, N. (2022). Forecasting Supply Chain Demand Approach Using Knowledge Management Processes and Supervised Learning Techniques. *International Journal of Information Systems and Supply Chain Management*, *15*(1), 1–21. doi:10.4018/IJISSCM.2022010103

Brahami, M., Atmani, B., & Matta, N. (2013). Dynamic Knowledge Mapping guided by Data Mining: Application on Healthcare. *The Journal of Information Processing Systems*, *9*(1), 1–31. doi:10.3745/JIPS.2013.9.1.001

Brahami, M., & Matta, N. (2019). The Relationship between the Knowledge Mapping and Open Innovation Process: The Case of Education System. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, *33*, 1–13. doi:10.1017/S0890060419000325

Brahami, M., Sabri, M., & Matta, N. (2020). Towards a Model to Improve Boolean Knowledge Mapping by Using Text Mining and Its Applications: Case Study in Healthcare. *International Journal of Information Retrieval Research*, *10*(3), 35–65. doi:10.4018/IJIRR.2020070103

Chakraborty, K., Bhatia, S., Bhattacharyya, S., Platos, J., Bag, R., & Hassanien, A. E. (2020). Sentiment Analysis of COVID-19 tweets by Deep Learning Classifiers—A study to show how Popularity is Affecting Accuracy in Social Media. *Applied Soft Computing*, *97*(106754), 1–15. doi:10.1016/j.asoc.2020.106754 PMID:33013254

Chen, K., Zhang, Z., Long, J., & Zhang, H. (2016). Turning from TF-IDF to TF-IGM for Term Weighting in Text Classification. *Expert Systems with Applications*, *66*, 245–260. doi:10.1016/j.eswa.2016.09.009

Daler, A., Saad Missen, M. M., & Husnain, M. (2021). Multiclass Event Classification from Text. *Scientific Programming (Hindawi)*, *2021*(660651), 1–15. doi:10.1155/2021/6660651

Desai, M., & Mehta, M. A. (2016). A Hybrid Classification Algorithm To Classify Engineering Students' Problems And Perks. *International Journal of Data Mining & Knowledge Management Process*, *6*(2), 73–81. doi:10.5121/ijdkp.2016.6206

Dhakate, P., Patil, S., Rajeswari, K., & Abin, D. (2014). Preprocessing and Classification in WEKA Using Different Classifiers. *International Journal of Engineering Research and Applications*, *4*(8), 91–93.

Divya, Y. A. (2018). Opinion Based Learning Model in Medical Sector. *International Journal of Scientific Research in Computer Science*. *Engineering and Information Technology*, *4*(6), 607–609.

Emadi, M., & Rahgozar, M. (2019). Twitter Sentiment Analysis using Fuzzy Integral Classifier Fusion. *Journal of Information Science*, 1–17.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). Knowledge Discovery and Data Mining: Towards a Unifying Framework. In *Proceeding of the 2nd International Conference on KDD & DM*. AAAI Press.

Foucault, N., & Courtin, A. (2016). Automatic Classification of Tweets for Analyzing Communication Behavior of Museums. *Proceeding of Tenth International Conference on Language Resources and Evaluation (LREC'2016)*.

Gautam, G., & Yadav, D. (2014). Sentiment analysis of Twitter Data using Machine Learning Approaches and Semantic Analysis. *Proceeding of the 7th International Conference on Contemporary Computing*, 437–442. doi:10.1109/IC3.2014.6897213

Go, A., Bhayani, R., & Huang, L. (2009). Twitter Sentiment Classification using Distant Supervision Processing. *Journal on Processing*, *1*(12), 1–6.

Gohil, S., Vuik, S., & Darzi, A. (2018). Sentiment Analysis of Health Care Tweets: Review of the Methods Used. *JMIR Public Health and Surveillance*, *4*(2), 1–10. doi:10.2196/publichealth.5789 PMID:29685871

Gridach, M., Haddad, H., & Mulki, H. (2018). Empirical Evaluation of Word Representations on Arabic Sentiment Analysis. *Communications in Computer and Information Science*, *782*, 147–158. doi:10.1007/978-3-319-73500-9_11

Hassan, S.-U., Saleem, A., Soroya, S. H., Safder, I., Iqbal, S., Jamil, S., & Nawaz, R. et al. (2020). Sentiment Analysis of Tweets Through Altmetrics: A Machine Learning Approach. *Journal of Information Science*, *2020*, 1–15.

Jalali, S. M. J., Moro, S., Mahmoudi, M. R., Ghaffary, K. A., Maleki, M., & Alidoostan, A. (2017). A Comparative Analysis of Classifiers in Cancer Prediction using Multiple Data Mining Techniques. *International Journal Business Intelligence and Systems Engineering*, *1*(2), 166–178. doi:10.1504/IJBISE.2017.088687

Kamakshi, P. (2020). Sentiment Analysis on Healthcare Tweets. *Indian Journal of Public Health Research & Development*, *11*(6), 566–569.

Kaoud, M. (2017). Investigation of Customer Knowledge Management: A Case Study Research. *International Journal of Service Science, Management, Engineering, and Technology*, *8*(2), 12–22. doi:10.4018/IJSSMET.2017040102

Kasture, N. R., & Bhilare, P. B. (2015). An Approach for Sentiment Analysis on Social Networking Sites. *Proceeding of the First International Conference on Computing, Communication, Control and Automation (ICCUBEA-2015)*, 390–395. doi:10.1109/ICCUBEA.2015.82

Kaur, J., Sehra, S. S., & Sehra, S. K. (2016). Sentiment Analysis of Twitter Data using Hybrid Method of Support Vector Machine and Ant Colony Optimization. *International Journal of Computer Science and Information Security*, *14*(7), 222–225.

Kaur, P., & Manik, S. (2018). Analysis of Data Mining and Soft Computing Techniques in Prospecting Diabetes Disorder in Human Beings: A Review. *International Journal of Pharmaceutical Sciences and Research*, *9*(7), 2700–2719.

Kaur, R., & Kautish, S. (2019). Multimodal Sentiment Analysis: A Survey and Comparison. *International Journal of Service Science, Management, Engineering, and Technology*, *10*(2), 38–58. doi:10.4018/IJSSMET.2019040103

Khalifa, N. M., Elghany, M. M., & Elghany, M. M. (2021). The Potential of Social Media in Emerging Supply Chain Management. *International Journal of Service Science, Management, Engineering, and Technology*, *12*(4), 39–58. doi:10.4018/IJSSMET.2021070103

Kumar, A., & Jaiswal, A. (2017). Empirical Study of Twitter and Tumblr for Sentiment Analysis using Soft Computing Techniques. *Proceedings of the World Congress on Engineering and Computer Science (WCECS'2017)*.

Madhuri, D. K. (2019). A Machine Learning based Framework for Sentiment Classification: Indian Railways Case Study. *International Journal of Innovative Technology and Exploring Engineering*, *8*(4), 441–445.

Mountassir, A., Benbrahim, H., & Berraba, I. (2013). Sentiment Classification on Arabic Corpora: A Preliminary cross-study. *Document Numérique, 16*(1), 73–96.

Munikar, M., Shakya, S., & Shrestha, A. (2019). Fine-grained Sentiment Classification using BERT. *Proceeding of the IEEE International Conference on Artificial Intelligence for Transforming Business and Society*, 1–4.

Nass, L., Swift, S., & Al Dallal, A. (2019). Indepth Analysis of Medical Dataset Mining: A Comparative Analysis on a Diabetes Dataset Before and After Preprocessing. *Annual PwR Doctoral Symposium 2018–2019, KnE Social Sciences*, 45–63.

Neethu, M. S., & Rajasree, R. (2013). Sentiment Analysis in Twitter using Machine Learning Techniques. In *Proceeding of the 4th International Conference on Computing, Communications and Networking Technologies (ICCCNT'2013)*. IEEE. doi:10.1109/ICCCNT.2013.6726818

O'leary, D. E. (2015). Twitter Mining for Discovery, Prediction and Causality: Applications and Methodologies. *Intelligent Systems in Accounting, Finance & Management*, *22*(3), 227–247. doi:10.1002/isaf.1376

Othman, E., & Al-Hamad, A. (2018). Automatic Arabic Document Classification Based on the HRWiTD Algorithm. *Journal of Software Engineering and Applications*, *11*(4), 167–179. doi:10.4236/jsea.2018.114011

Özçift, A. (2020). Medical Sentiment Analysis based on Soft Voting Ensemble Algorithm. *Journal of Yönetim Bilişim Sistemleri Dergisi*, *6*(1), 42–50.

Pang, B., & Lee, L. (2004). A Sentimental Education: Sentiment Analysis using Subjectivity Summarization based on Minimum Cuts. *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics (ACL - 04)*, 271–278.

Picassoa, A., Merelloa, S., Ma, Y., Onetoa, L., & Cambria, E. (2019). Technical Analysis and Sentiment Embeddings for Market Trend Prediction. *Expert Systems with Applications*, *135*, 60–70. doi:10.1016/j.eswa.2019.06.014

Pratibha, D. S., Ragin, O. R., & Anisha, G. S. (2019). Comparative Study of Classification Algorithms in Chronic Kidney Disease. *International Journal of Recent Technology and Engineering*, *8*(1), 180–184.

Putra, N. P., & Maulany, G. J. (2018). Classification System for Student Study Duration on Department of Information Systems at Musamus University Using ID3. *International Journal of Mechanical Engineering and Technology*, *9*(12), 878–885.

Rahab, H., Zitouni, A., & Djoudi, M. (2019). SANA: Sentiment Analysis on Newspapers Comments in Algeria. *Journal of King Saud University – Computer and Information Sciences, 31*(2), 1–9.

Rathi, M., Malik, A., Varshney, D., Sharma, R., & Mendiratta, S. (2018). Sentiment Analysis of Tweets Using Machine Learning Approach. *Proceedings of Eleventh International Conference on Contemporary Computing (IC3¢2018)*. doi:10.1109/IC3.2018.8530517

Ravi, A., Khettry, A. R., & Yelandur Sethumadhavachar, S. (2019). Amazon Reviews as Corpus for Sentiment Analysis Using Machine Learning. *Advances in Computing and Data Sciences*, 403–411.

Ridgeway, G. (2002). Looking for Lumps: Boosting and Bagging for Density Estimation. *Computational Statistics & Data Analysis*, *38*(4), 379–392. doi:10.1016/S0167-9473(01)00066-4

Sanjudevi, R., & Savitha, D. (2019). Dengue Fever Prediction Using Classification Techniques. *International Research Journal of Engineering and Technology*, *6*(2), 558–563.

Sankaranarayanan, J., Samet, H., Teitler, B. E., Lieberman, M. D., & Sperling, J. (2009). TwitterStand: news in tweets. *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 42–51.

Sewalk, K. C., Tuli, G., Hswen, Y., Brownstein, J. S., & Hawkins, J. B. (2018). Using Twitter to Examine Web-Based Patient Experience Sentiments in the United States: Longitudinal Study. *Journal of Medical Internet Research*, *20*(10), 1–14. doi:10.2196/10043 PMID:30314959

Sharma, P., & Kaur, M. (2013). Classification in Pattern Recognition: A Review. *International Journal of Advanced Research in Computer Science and Software Engineering*, *3*(4), 298–306.

Singaravelan, S., Murugan, D., & Mayakrishnan, R. (2015). Analysis of Classification Algorithms J48 and SMO on Different Datasets. *World Engineering & Applied Sciences Journal*, *6*(2), 119–123.

Singhvi, V., & Srivastava, P. (2021). Evaluation of Consumer Reviews for Adidas Sports Brands Using Data Mining Tools and Twitter APIs. *International Journal of Service Science, Management, Engineering, and Technology*, *12*(6), 89–104. doi:10.4018/IJSSMET.2021110106

Sirsat, S., Rao, S., & Wukkadada, B. (2019). Sentiment Analysis on Twitter Data for Product Evaluation. *Proceedings of the 5th Somaiya International Conference On Technology And Information Management (SICTIM'2019).*

Spencer, J., & Uchyigit, G. (2012). Sentimentor: Sentiment Analysis on Twitter Data. *Proceedings of the 1st International Workshop on Sentiment Discovery from Affective Data*, 65–66.

Srivastava, A., Singh, V., & Singh Drall, G. (2019). Sentiment Analysis of Twitter Data: A Hybrid Approach. *International Journal of Healthcare Information Systems and Informatics*, *14*(2), 1–16. doi:10.4018/IJHISI.2019040101

Swathi, H., Aravinth, S. S., Nivethitha, V., Saranya, T., & Nivethanandhini, R. (2019). Sentiment Analysis of Movie Review using Data Analytics Techniques. *Iconic Research and Engineering Journals*, *2*(9), 48–52.

Tanulia, Y., & Girsan, A. S. (2019). Sentiment Analysis on Twitter for Predicting Stock Exchange Movement. *Advances in Science*. *Technology and Engineering Systems Journal*, *4*(3), 244–250. doi:10.25046/aj040332

Tariyal, A., Goyal, S., & Tantububay, N. (2018). Sentiment Analysis of Tweets Using Various Machine Learning Techniques. *Proceedings of the Eleventh International Conference on Contemporary Computing (IC3¢2018)*. doi:10.1109/ICACAT.2018.8933612

Ur-Rahman, N., & Harding, J. A. (2012). Textual Data Mining for Industrial Knowledge Management and Text Classification: A Business Oriented Approach. *Expert Systems with Applications*, *39*(5), 4729–4739. doi:10.1016/j.eswa.2011.09.124

Velammal, B. L. (2019). Development of Knowledge based Sentiment Analysis System using Lexicon Approach on Twitter Data. *International Journal of Knowledge Management Studies*, *10*(1), 58–69. doi:10.1504/IJKMS.2019.097125

Vijiyarani, S., & Sudha, S. (2013). Disease Prediction in Data Mining Technique–A survey. *International Journal of Computer Applications & Information Technology*, *2*(1), 17–21.

Wahi, A. K., Medury, Y., & Misra, R. K. (2014). Social Media: The Core of Enterprise 2.0. *International Journal of Service Science, Management, Engineering, and Technology*, *5*(3), 1–15. doi:10.4018/ijssmet.2014070101

Wang, N., Varghese, B., & Donnelly, P. D. (2016). A Machine Learning Analysis of Twitter Sentiment to the Sandy Hook Shootings. *Proceedings of the IEEE 12th International Conference on eScience*. doi:10.1109/eScience.2016.7870913

Wongsirichot, T., Kajkamhaeng, S., Nupinit, W., & Sridonthong, N. (2019). An Investigation of Data Mining Based Automatic Sleep Stage Classification Techniques. *International Journal of Machine Learning and Computing*, *9*(4), 520–526. doi:10.18178/ijmlc.2019.9.4.835

Yasen, M., & Tedmori, S. (2019). Movies Reviews Sentiment Analysis and Classification. *Proceedings of the IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)*, 860–865.

Yelmen, I., Zontul, M., Kaynar, O., & Sonmez, F. (2018). A Novel Hybrid Approach for Sentiment Classification of Turkish Tweets for GSM Operators. *International Journal of Circuits*, *Systems and Signal Processing*, *12*, 637–645.

Zafar, S., Kamran, M., & Hu, X. (2019). Intrusion-Miner: A Hybrid Classifier for Intrusion Detection using Data Mining. *International Journal of Advanced Computer Science and Applications*, *10*(4), 329–336. doi:10.14569/IJACSA.2019.0100440

Zaidan, A. A., Zaidan, B. B., Kadhem, Z., Larbani, M., Lakulu, M. B., & Hashim, M. (2015). Challenges, Alternatives, and path to Sustainability: Better public Health Promotion using Social Networking Pages as a Key Tool. *Journal of Medical Systems*, *39*(2), 2–14. doi:10.1007/s10916-015-0201-y PMID:25600194

Zhao, L., Chen, J., Chen, F., Wang, W., Lu, C., & Tech, V. (2015). SimNest: Social Media Nested Epidemic Simulation via Online Semi-Supervised Deep Learning. *Proceedings of the IEEE International Conference on Data Mining (ICDM'15),* 639–648.

Brahami Menaouer is an associate professor of computer science in the Systems Engineering department at National Polytechnic School of Oran (Algeria). He received a HDR degree in computer science from the University of Science and Technology of Oran Mohamed Boudiaf (Algeria), in 2018. He received a Ph.D. in computer science from the University of Oran (Algeria). His research interests include Knowledge Management, Knowledge Mapping, Project Management, Business Intelligence, Data Mining, Machine Learning, Deep Learning and Boolean Modeling. His work has appeared in Extraction and Knowledge management, Journal of Information Processing Systems, International Journal of Computer Science Issues, Journal of Computer Applications, International Journal of Information Systems in the Service Sector, Procedia - Social and Behavioral Sciences (Elsevier), European Journal of Social Law, Artificial Intelligence for Engineering Design, Analysis and Manufacturing, and International Journal of Healthcare Information Systems and Informatics.

Fatma Abdeldjouad received her first Degree in Computing at National Polytechnic School of Oran, Systems Engineering department, Algeria. She furthered her master study in the area of Knowledge Management and Artificial Intelligence in 2018. Her research interests are knowledge management, machine learning and healthcare studies to improve her work.

Sabri Mohammed is a research professor of computer science in the National Polytechnic School of Oran (Algeria). He received a Ph.D. in computer science from the University of Science and Technology of Oran Mohamed Boudiaf (Algeria), in 2018. His research interests include Service Oriented Architecture (SOA), Knowledge Representation, Knowledge Visualization, Data Mining, and Business Intelligence. His work has appeared in International Journal of Information Technology and Computer Science.