



Article

An Intelligent Hybrid Sentiment Analyzer for Personal Protective Medical Equipments Based on Word Embedding Technique: The COVID-19 Era

Ruba Obiedat ¹, Laila Al-Qaisi ², Raneem Qaddoura ³, Osama Harfoushi ¹ and Ala' M. Al-Zoubi ^{1,4,*}

- ¹ King Abdullah II School for Information Technology, The University of Jordan, Amman 11942, Jordan; r.obiedat@ju.edu.jo (R.O.); o.harfoushi@ju.edu.jo (O.H.)
- ² Information Systems and Networks Department, Faculty of Information Technology, The World Islamic Sciences and Education University, Amman 11947, Jordan; Laila.qaisi@wise.edu.jo
- ³ School of Computing and Informatics, Al Hussein Technical University, Amman 11831, Jordan; raneem.qaddoura@htu.edu.jo
- ⁴ School of Science, Technology and Engineering, University of Granada, 18010 Granada, Spain
- * Correspondence: alzoubi@correo.ugr.es or ala.m.zoubi@gmail.com

Abstract: Due to the accelerated growth of symmetrical sentiment data across different platforms, experimenting with different sentiment analysis (SA) techniques allows for better decision-making and strategic planning for different sectors. Specifically, the emergence of COVID-19 has enriched the data of people's opinions and feelings about medical products. In this paper, we analyze people's sentiments about the products of a well-known e-commerce website named Alibaba.com. People's sentiments are experimented with using a novel evolutionary approach by applying advanced pre-trained word embedding for word presentations and combining them with an evolutionary feature selection mechanism to classify these opinions into different levels of ratings. The proposed approach is based on harmony search algorithm and different classification techniques including random forest, k-nearest neighbor, AdaBoost, bagging, SVM, and REPTree to achieve competitive results with the least possible features. The experiments are conducted on five different datasets including medical gloves, hand sanitizer, medical oxygen, face masks, and a combination of all these datasets. The results show that the harmony search algorithm successfully reduced the number of features by 94.25%, 89.5%, 89.25%, 92.5%, and 84.25% for the medical glove, hand sanitizer, medical oxygen, face masks, and whole datasets, respectively, while keeping a competitive performance in terms of accuracy and root mean square error (RMSE) for the classification techniques and decreasing the computational time required for classification.

Keywords: word embedding; sentiment; medical products; metaheuristic; pre-trained models



Citation: Obiedat, R.; Al-Qaisi, L.; Qaddoura, R.; Harfoushi, O.; Al-Zoubi, A.M. An Intelligent Hybrid Sentiment Analyzer for Personal Protective Medical Equipments Based on Word Embedding Technique: The COVID-19 Era. *Symmetry* **2021**, *13*, 2287. <https://doi.org/10.3390/sym13122287>

Academic Editors: Sergei D. Odintsov and Carlo Cattani

Received: 22 October 2021
Accepted: 27 November 2021
Published: 1 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The use of social media is growing rapidly; it represents now, more than ever, a big part of our lives [1,2]. People from different countries are sharing their feelings, thoughts, attitudes, opinions, and concerns about different life aspects on these platforms on daily basis. The massive amount of data gained by different social channels has become the main source of information for different domains such as business, governments, and health [3]. Thus, the fast growth of information combined with the existence of advanced data mining and sentiment analysis (SA) techniques present an opportunity to mine these information in different sectors [4]. Analyzing these data is essential for decision-making and strategic planning in these fields [3,5].

Medhat et al. [6] has defined sentiment analysis (SA) or opinion mining (OM) as “the computational study of people's opinions, attitudes and emotions toward an entity”. Sentiment analysis is a multidisciplinary field that focuses on analyzing people attitudes, reviews, feedback, and concerns toward different aspects of life including products, services,

companies, and politics, using different techniques such as natural language processing (NLP), text mining, computational linguistics, machine learning, and artificial intelligence for enhancing the decision making process [7].

Sentiment analysis can be applied on three different classification levels: document level, sentence level, and aspect level [6]. It has three main approaches; the machine learning approach, lexicon-based approach, and hybrid approach. Machine learning (ML) relies on the main ML techniques with the use of syntactic and linguistic features [6,8]. Many ML techniques can be applied for sentiment analysis including support vector machine (SVM), decision tree (DT), K-nearest neighbors (k-NN), naïve Bayes (NB), and others [9]. Gautam and Yadav [10], for example, followed a machine learning approach to analyze people's sentiment about certain products on Twitter. First, consumers' reviews about the product were collected; then, different machine learning algorithms were applied, including SVM, naïve Bayes, and maximum entropy; and finally, the performance of the different classifiers was measured. In addition, Samal et al. [11] used a machine learning approach to perform sentiment analysis on a movie review dataset. The study collected the dataset and then used common supervised machine learning algorithms, namely naïve Bayes, multinomial naïve Bayes, Bernoulli naïve Bayes, logistic regression, stochastic gradient descent (SGD) classifier, linear SVM/ linear SVC, and Nu SVM/Nu SVC, to train the model, find its accuracy, and then compare the results. In this paper, a machine learning (ML) approach is applied by following a supervised classification technique to analyze peoples' sentiments about the medical products during the COVID-19 pandemic.

Sentiment analysis can be applied in several life applications and employed in different sectors from marketing and finance to health and politics [12]. People are expressing their feelings about products, services, events, organizations or public figures [13].

In politics, for example, candidates in different government positions can use sentiment analysis in running their campaigns and sense the feeling of voters toward different country issues or different actions of the candidate [14]. Hasan et al. [15], for instance, has followed a machine learning approach to perform sentiment analysis on political views during elections. The work collects reviews from the Twitter platform and then applies naïve Bayes and support vector machines (SVM) to test the accuracy of the results.

Another important domain employing sentiment analysis is during disease epidemics and natural disasters. Sentiment mining can determine how people react during these disasters and how to use this information in managing these disasters in a better manner. SA was ranked as the fourth main source of information during emergencies; people can post their experience in text, photos, or videos, express their feelings, panics, and concerns, and report problems, make donations, and express support for authorities. Analyzing these sentiments helps improve the management of the epidemic [7].

Customers' reviews of products and services present the most common application of sentiment analysis. Thus, consumer sentiment analysis (CSA) has become a trend recently [14]. There are many websites that specialize in summarizing customers' reviews about certain products, such as "Google Product Search" [14]. Since there is a massive amount of available reviews, filtering the most relevant reviews and analyzing them not only speeds up the decision process but also improves its performance [16]. SA of customers' reviews can be used to enhance the business values, ensure customer loyalty and raise the quality of products and services. All that will be reflected in the company's image and customers' satisfaction, which leads to more sales and higher revenues [9,17].

Sentiment analysis in the medical domain is not broadly applied yet [18], although a study by the Pew Internet and American Life Project showed that 80% of internet users are visiting health-related topic online, 63% of people searching for information about certain medical problems, while 47% are looking for medical treatment and products online [19].

Rozenblum and Bates [20] has described the social web and the internet in terms of patient-centered health care as a "perfect storm", because it presents a valued source of information for the public and health organizations, since patients increasingly describe, share, and rate their experience of medical products and services over the internet. With

this huge amount of data, it is essential to collect and analyze this information by capturing the medical sentiment, which will be helpful for patients, decision-makers, and the whole health sector [19].

In this context, Jiménez-Zafra et al. [4] analyzed people's opinions posted in medical forums regarding doctors and drugs. It was found that drug reviews were more difficult to analyze than doctors' reviews. Although both reviews were written in informal language by non-professional users, drug reviews have greater lexical diversity. Abualigah et al. [21] had presented a brief review about using sentiment analysis in analyzing data about people's experience in healthcare medication, treatments, or diagnosis posted on their personal blogs, online forums, or medical websites. Patients visited different healthcare centers and shared their experiences concerning services, pleasure, and availability. Using sentiment analysis can help patients learn from others' experiences, report medical problems and resolve them, as well as improve medical decisions and increase healthcare quality. Another recent study by Polisená et al. [22] performed a scoping review about using sentiment analysis in health technology assessment (HTA). The study used the patients' posts on different social media platforms about the effectiveness and safety of these health technologies such as medical devices, HPV vaccination, and drug therapies.

The COVID-19 outbreak began in late December 2019 and spread rapidly worldwide. The World Health Organization (WHO) announced it as a global pandemic on 11 March 2020 [23]. The pandemic had a negative effect on the biggest companies in different sectors and the whole productive system around the globe [12]. On the other hand, some companies in the medical product sector producing personal protective equipment (PPE) witnessed an increased demand for their products during the pandemic. The World Health Organization (WHO) named 17 products as the key products needed to deal with the pandemic [24], including personal protection equipment such as gloves and face masks or some medical devices for case management, such as oxygen sensors, oxygen concentrators, and respirators, in addition to sterilizers and pharmaceutical companies [25]. These medical supplies have suffered from a dramatic shortage during the pandemic due to their huge demand around the globe [24,25].

Due to the extreme importance of PPE products during the pandemic and the availability of different brands with different qualities on social commerce websites, people are seeking the best product and searching for customers' reviews about these medical products online. Even before the pandemic, people were obtaining information about different products from various social media platforms. Nowadays, when any person wants to buy a new product online, they will search for people's reviews and comments about that product on different social channels [14]. A study conducted by Deloitte has a similar opinion; it stated that "82% of purchase decisions have been directly influenced by reviews" [26].

This research studies people's sentiment regarding the main personal protective equipment (PPE) used during the pandemic including; face masks, medical gloves, hand sanitizer, and medical oxygen. Face masks, such as surgical masks and N95 respirators, represent the most essential product in fighting COVID-19. These masks are used by patients, healthcare workers, and all people to protect them from being infected. Certain types of masks can be reused or worn for a longer duration under certain circumstances. It was found that the demand for facial masks had increased ten times during the pandemic compared to world production prior to the crisis, and the price has increased as well [27]. In addition, medical gloves have a similar usage during the COVID-19 outbreak, especially when contacting any potentially infectious patients or materials. People become more aware of the importance of using gloves, and it is expected that they will continue using them even after the epidemic [28]. Hand sanitizer is another important product spread widely during the pandemic; it can be found in the entrances of hospitals, companies, malls, shops, schools, and all different buildings. People are buying and using it extensively and regularly [25]. Furthermore, in the context of the current pandemic, medical oxygen presents an essential step for the recovery of patients with severe COVID-19 [29]. According

to Stein et al. [29], 41.3% of COVID-19 hospitalised patients in China need supplemental oxygen. Thus, huge supplies of medical oxygen are needed during the pandemic, hospitals and ICUs are at full capacity, and the whole health system is overwhelmed [30]. Therefore, home oxygen concentrators have become one of the primary sources of oxygen during the pandemic, since patients with middle-to-severe symptoms need home monitoring and treatment [31].

Motivated by the importance of the aforementioned medical products in saving lives during the pandemic and the significance of analyzing people sentiments regarding these products, this study takes the initiative to collect people's comments and reviews about COVID-19 medical products from Alibaba.com, one of the biggest e-commerce websites in the world to find the best PPE with the highest quality.

The study proposes a novel evolutionary approach aiming to classify people's sentiment towards these products, as online reviews are the most influential factor in the decision-making process. Moreover, the following approach applies an evolutionary feature selection method to select the best subset of features and enhance the classification performance in terms of accuracy and computational time. On the other hand, the proposed approach can be used in handling such crisis and medical situations quickly and efficiently in the future.

Furthermore, the contribution of this study can be summarized by the following points:

- A new sentiment analysis study for classifying people's opinions towards medical products that are related to the COVID-19 crisis, including gloves, hand sanitizers, face masks, and home oxygen concentrators, to provide decision-makers with analyzed observations of customers' feedback to help them take prompt actions of the effectiveness of the products.
- Applying advanced pre-trained word embedding learning techniques for feature extraction and word presentation to overcome the challenges of the data.
- Conduct an evolutionary feature selection method to select the best subset of features, which are extracted by the word embedding technique, using different classifiers for evaluation.

The remainder of the paper is divided as follows: In Section 2, a literature review on sentiment analyses during COVID-19 is performed. A brief overview of the methods and concepts is described in Section 3. Section 4 describes the proposed approach. Section 5 explains the experiments and results of this study. Finally, Section 6 summarizes the conclusions future directions.

2. Related Work

Sentiment analysis (SA) can be briefly defined as detecting feelings expressed implicitly within the text. This text may be describing opinion, review, or behaviour towards an event, product, or organization [32]. SA has been a trend during the last decade to facilitate decision-making and support domain analysts' missions in a wide range of practical applications such as healthcare, finance [33], media, consumer markets [34], and government [35,36].

Due to the widespread use of online platforms such as Twitter, Facebook, Instagram, and WhatsApp, especially during the COVID-19 lockdown internationally, people are posting online reviews regarding used products. This resulted in extensive usage of consumer sentiment analysis (CSA) using online reviews [37]. These online reviews in which people indicate their opinions or attitude toward a service or product are found to be beneficial to organizations in analyzing customers' experience for improvement purposes [38].

In the medical domain, many research papers study the reviews of customers in various aspects. Gohil et al. [39] proposed a systematic literature review (SLR) on healthcare sentiment analysis by reviewing research studies targeting hospitals and healthcare reviews published on Twitter, while Lagu et al. [40] studied patient online reviews and

their experience expressed using physician-rating websites in the United States. There were 66 potential websites, 28 of which met the inclusion criteria set of the study, with 8133 reviews found for 600 physicians. Additionally, Liu et al. [41] examined patients' satisfaction towards online pharmaceutical websites. The study involved several aspects of overall online B2C aspects, namely product, logistics, factors, price, information, and system used. Moreover, the Na and Kyaing [42] proposed a SA method based on the lexicon, grammatical relations, and semantic annotation on drug review websites. The approach was examined on 2700 collected reviews, and results found that applying it would be useful to drug makers and clinicians as well as patients.

SA is conducted based on the lexicon, machine learning (ML), or graph [35]. Lexicon-based works can be found in some studies. Jiménez-Zafra et al. [4] examined combining supervised learning and lexicon-based SA for reviews about both drugs and doctors. Reviews were extracted from two Spanish forums: DOS for drug reviews and COPOS for reviews about physicians. SVM classifier were applied with four different word representations (TF-IDF, TF, BTO, and Word2Vec).

Many researchers conducted SA using ML in the medical field. Gräßer et al. [43] examined SA on pharmaceutical review sites. The aim was to predict the overall satisfaction, side effects, and effectiveness of specific drugs. Data were extracted from two web pages, Drugs.com and Druglib.com. The ML classifier considered in the study was logistic regression (LR). Results were found to be promising, and a further extension by applying deep learning (DL) was recommended. Moreover, Daniulaityte et al. [44] applied a supervised ML SA to classify tweets published about drug abuse. Data were extracted from Twitter through Twitter streaming application programming interface (API). Logistic regression (LR), naïve Bayes (NB), and support vector machines (SVM) classifiers were used with 5-fold cross-validation along with F1 measure as an evaluation metric. Another study performed by Basiri et al. [45] analyzed medical and healthcare reviews in which two novel deep fusion frameworks were proposed. The first, (3W1DT), used the DL model as a base classifier (BC) and a traditional ML classifier such as naïve Bayes (NB) or decision tree (DT) as a secondary classifier (SC). The second, 3W3DT, used three DL and one traditional ML model. In both DL models, ML models including NB, DT, RF, and KNN were combined to improve classification performance. Data were extracted from Drugs.com, and proposed models were tested in terms of accuracy and F1-measure.

Other researchers combined lexicon-based with ML techniques. A study by Harrison and Sidey-Gibbons [46] conducted a natural language processing (NLP)-based SA on the following drugs: Levothyroxine, Viagra, Oseltamivir, and Apixaban in three main phases. The first phase applied lexicon-based SA, and the second one used unsupervised ML (latent Dirichlet allocation, LDA) to distinguish reviews written on similar drugs within the dataset. The third phase considered predicting negative or positive reviews on drugs using three supervised ML algorithms: regularised logistic regression, a support vector machine (SVM), and an artificial neural network (ANN). Finally, a comparison analysis was performed between the used classifiers based on classification accuracy, the area under the receiver operating characteristic curve (AUC), sensitivity, and specificity.

However, lexicon-based SA suffers from several drawbacks, especially in the medical field. Firstly, it relies on a predefined list of polarity annotation, which differs based on the used language. Secondly, the language used in reviews is usually informal or slang, which probably is not included in the lexicon. Thirdly, medical terms' and words' meanings may differ depending on the context, which may be misleading. As a result, ML-based SA was found to provide a good alternative [39].

In addition, the aforementioned studies that rely on using ML techniques for medical applications considered sentiments for the reviews of drugs and doctors. They have not considered analyzing sentiments about medical products for COVID-19 crises, which allows decision-makers to take prospective and immediate actions toward improving these products during the crisis.

On the other hand, as stated by Karyotaki et al. [47], a reviewer's positive mood is affected by his attention to any internal or external condition, that is, in our domain, purchasing, reviewing comments, or any nearby customer experience denoted in word of mouth (WOM), while [48,49] proved that emotional intelligence has a significant impact on the research area, which should be taken into consideration, and emphasized the importance of studying both feelings and emotions within a specific domain.

Moreover, extracting new medical product reviews and applying pre-trained word embedding for sentiment analysis is new, since applying pre-trained word embedding for sentiment medical products and using wrapper fs for multiple classifiers at the same time for medical sentiment analysis, where each classifier is used as an evaluation for the wrapper FS, are not present in other studies. In addition, studying the quality of the medical products during COVID-19 is very important in the meantime.

As a result, this paper aims to analyze opinions and feelings expressed through Alibaba online e-commerce websites towards medical products, including medical gloves, hand sanitizer, face masks, and whole datasets, which are extensively consumed during the pandemic condition of COVID-19.

3. Preliminaries

3.1. Ordinal Regression

Originally, ordinal regression as a term is a statistical regression analysis and might be referred to as ordinal classification. It aims to predict ordinal data, that is, statistical data with natural and categorized variables without knowing distances between categories. Ordinal regression problems relay between classification and regression. It has been used in other fields such as psychology and social sciences. It is worth mentioning here that modeling human preference levels (on a scale from, for example, 1–5 for “very poor” through “excellent”) presents an example of ordinal regression [50].

In ML, ordinal regression can be referred to as ranking learning, aiming to obtain patterns classified into naturally ordered labels based on a categorical scale. An ordinal regression classifier performs as a ranking function, in which class labels are interpreted into scores. This process results in obtaining a huge amount of scores and ties. This may be a drawback for ranking techniques such as multipartite ranking. Yet, it can be solved by including thresholds for setting intervals for every single class [51].

3.2. Evolutionary Algorithm

Evolutionary algorithms (EA) are classified under biologically inspired algorithms and based on Darwinian evolution theory and usually exploit a set of various synthetic methods such as population management, replication, variability, and selection. They were found to be simple at high-level problems but became complex when knowledge is extracted from the problem domain [52]. They have powerful capabilities in solving machine learning problems, which is a growing area of research nowadays [53].

They are also referred to as metaheuristics, which can be defined as higher-order algorithms that aim to systematically identify the best solution within the problem space. EA algorithms are in this category as well [54]. EA algorithms are executed in order to achieve an already set quantitative goal such as time function. In addition, this goal is a metric of success and an exit method for the algorithm itself, which may be either a single-objective or multi-objective goal. They are known for their effectiveness in finding optimal solutions for optimization problems [55].

The most popular and studied examples are particle swarm optimizers (PSO), genetic algorithms (GA), ant colony optimization (ACO), grasshopper optimization algorithm (GOA), and harmony search algorithm (HSA).

3.2.1. Harmony Search Algorithm

HSA is a meta-heuristic optimization algorithm that was firstly proposed in 2001; it is based on designing a unique harmony in music for addressing optimization problems.

The main steps of HSA are clarified by the pseudo-code in Algorithm 1, in which every single possible solution x is necessary $x \in X$ depending on certain N decision variables, and X represents the space of whole possible solutions. The main target is to find either a minimum or maximum value of fitness function F_{X_i} for all $i = 1, \dots, N$. Pitch adjustment rate (PAR) and bandwidth (BW) are both used to control the exploitation process in HSA, while exploration is the responsibility of harmony memory considering rate (HMCR), as explained in [56].

HSA has five main steps, as explained by [56,57]. First is the initialization of basic HS parameters, HMCR, BW, PAR, number of iterations (NI), and HM size (HMS), as well as determination of the optimization problem goal. Second is the initialization of HM values, as X_i should be within upper and lower boundaries using Equation (1), where R denotes a random number between 0 and 1. Third is new harmony improvisation through combining HMCR, PAR, and BW. To do so, let $x_{i,j}$ be the value of the j th decision variable of i th solution; a random number is generated between 0 and 1, and improvisation will be achieved as in Equation (2).

$$X_i = LB_i + R \times (UB_i - LB_i) \quad (1)$$

$$x_{new,j} \leftarrow \begin{cases} x_{new,j} \in \{x_{1,j}, x_{2,j}, \dots, x_{S,j}\} & R < HMCR \\ x_{new,j} = x_{i,j} \pm R \times BW & R < PAR \\ x_{new,j} = LB_i + R \times (UB_i - LB_i) & \text{Otherwise} \end{cases} \quad (2)$$

Forth is a fitness function assessment for the newly produced solution; it is checked whether it is superior to the worst solution within HM; then, an updated memory is conducted by replacing the worst with a new solution. Finally, the fifth step is checking the stop criteria set; if it is met, the algorithm will be terminated.

Algorithm 1 Search Algorithm (HSA) pseudo-code [56]

- 1: **Input:** Generate the initial harmonies randomly
 - 2: **Output:** Optimal solution with its fitness value
 - 3: **Algorithm**
 - 4: Initialize the parameters of the HS *HMCR*, *PAR* and etc.
 - 5: Initialize harmony memory (HM)
 - 6: Evaluate all solutions using the fitness function
 - 7: **while** Termination criteria **do**
 - 8: new solution = ϕ
 - 9: **if** *rand* < *HMCR* **then**
 - 10: Memory consideration
 - 11: **if** *rand* < *PAR* **then**
 - 12: Pitch adjustment
 - 13: **else**
 - 14: Random consideration
 - 15: Evaluate the fitness function of the new solution
 - 16: Replace the worst solution in HM by the new solution
-

Its advantages may be summarized as execution clarity, ability to tackle complex problems, and registered success and lower mathematical processes required, as mentioned in [57].

3.2.2. Wrapper Feature Selection

Feature selection (FS) is the process of reducing symmetrical features used to build a model in data mining [58]. In other words, It may be referred to as the process of selecting a subset of the most relevant features to be used for data mining model construction. Mostly, FS is highly recommended and applied for avoiding the curse of dimensionality. There are three types of FS, which are filter, wrapper, and embedded methods [59].

Wrapper FS methods mainly depend on a certain ML algorithm applied to the dataset. It runs a greedy search approach to examine all possible subsets, which obtain the best value of an evaluation criterion. Different performance metrics are meant by evaluation criterion; this is chosen based on the problem domain and has very popular examples such as accuracy, precision, recall, and f1-score [60].

3.3. Word Embedding Feature Extraction

Feature extraction (FE) is one of the core steps within data mining projects. It is used to determine features that have a positive effect through classification [61]. Many techniques are used to accomplish this task, such as bag-of-words, term frequency–inverse document frequency (TF–IDF), word embedding (WE), and natural language processing (NLP)-based techniques such as word count, noun count, and word2vec [62].

Word embedding is a term used to refer to word representation for text analysis; other used terms may be word representation and distributed word representation [63,64]. Typically, word meanings are encoded as vectors of real values in which close values in vector space indicate similar words in meaning. Both language modeling and feature learning techniques are required to map words or phrases from vocabulary into real number vectors to obtain WE [65].

Word2vec Embeddings

Word2vec is a technique published in 2013 by a research team in Google led by Tomas Mikolov for NLP over two studies [66,67]. It is an artificial neural network (ANN)-based group of models that are used to build word embeddings through extracting word association from text within a corpus. After conducting the training phase, such a model can both detect synonyms and predict additional words for sentences. It represents every single word in form of vectors. A simple mathematical function is used to indicate semantic similarity between vectors. Nevertheless, it was mentioned tehabib2021altibbivec that Word2vec is a shallow and two-layer ANN, which takes text corpus as input and produces a vector space. Word2vec can use either of two main model architectures, namely continuous bag-of-words (CBOW) or continuous skip-gram (SG), as shown in Figure 1.

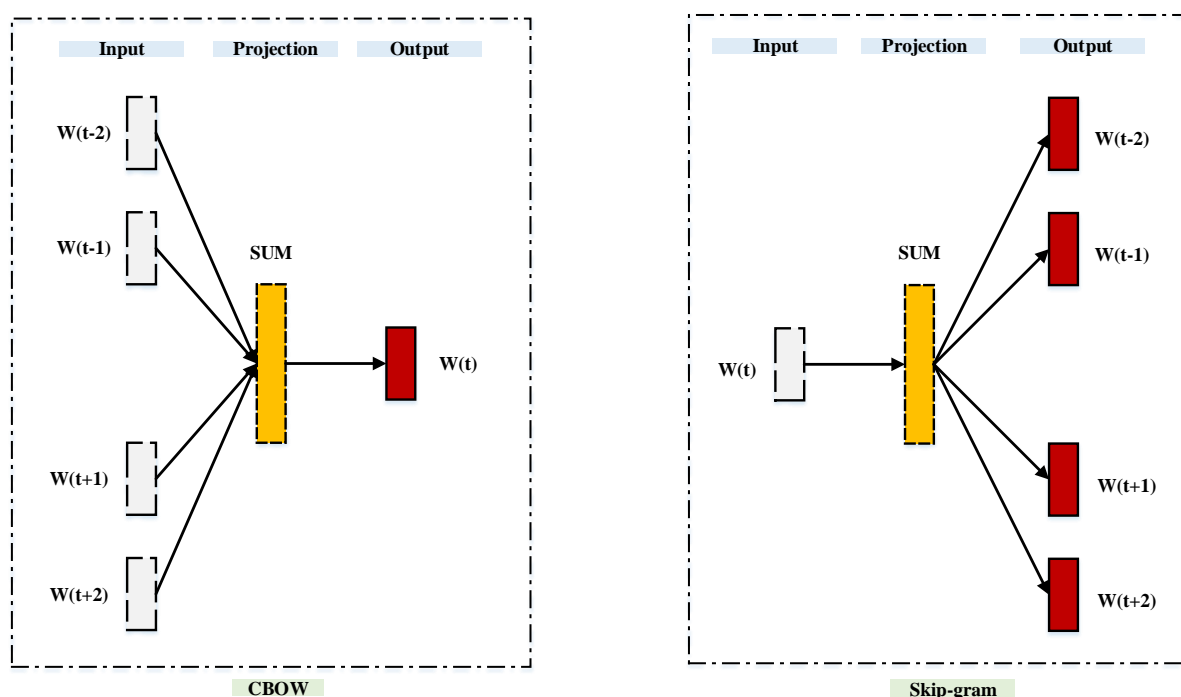


Figure 1. Word2Vec's CBOW and SG models.

4. Methodology

In this section, the methodology processes of the proposed work have been presented. These processes consist of data description and collection, data preparation, and proposed approach, each of which will be discussed in detail in the following subsections.

4.1. Data Description and Collection

This work investigates the feedback of Alibaba e-commerce website customers reviews regarding several medical products during the COVID-19 situation. Alibaba is a Chinese multinational e-commerce corporation for retail, internet, and technology. The website was created in 1990 in Hangzhou, specializing in business-to-business (B2B), consumer-to-consumer (C2C), and business-to-consumer (B2C) online sales services. The website also provides different services including cloud computing, electronic payment, and shopping search engines services. In the meantime, Alibaba is considered one of the biggest e-commerce and retailer corporations. Additionally, it is ranked as the fifth-biggest artificial intelligence firm in 2020.

One of the main reasons to select the Alibaba website is due to the availability of medical products as well as the existence of many consumers' reviews for these products compared with other symmetrical websites. In this work, we are aiming to enhance the quality of products through advanced sentiment analysis of the consumer reviews due to its importance and high need from various parties, especially at this time.

The collection process was performed using a crawler tool to gather the reviews of each product on the website. Each review consists of the consumer name, the context of the review, review date, and the rating of the review, either of 1, 2, 3, 4, or 5, according to the person's assessment. It is worth noting that the ratings were crawled as images of 1 to 5 stars.

The collected data consist of medical products such as gloves, hand sanitizers, face masks, and home oxygen concentrators. We choose these products as they were the most used commodities in the last two years.

4.2. Data Preparation

The preparation of the data took various phases to be ready for the experimentation step by the models. These pre-processing phases consist of removing missing values and stop words, and cleaning, normalization, and formatting the data. The reviews did not require to be labeled, since they had already been rated by the customer. Nevertheless, to ensure that the labeling was accurate, some experts were asked to read a sample of the data.

Therefore, after the preprocessing phases, the feature extraction is performed through several steps. First, we conduct a tokenization process to divide the text into a set of words. In other words, the text is transformed from raw textual data into several features, where each feature represents some kind of statistical measurement of that word. Various text vectorization models have been proposed in the literature. In terms of popularity, term frequency, term frequency-inverse document frequency, and bag-of-words are the most commonly used models. Despite their popularity, however, they do not reveal the actual meaning (semantic) of the words. Therefore, to understand the implicit relationships between words, a more comprehensive approach is needed, namely word embedding.

Embedding is the process of representing words as numerical vectors, in which the words are encoded as dense vectors that create a unique analogous encoding of all words of similar meaning as shown in Figure 2. During the training process, the density vector's components (parameters) are learned and determined. By increasing the embedding dimension, the learning ability is enhanced, although a much larger training dataset is required.

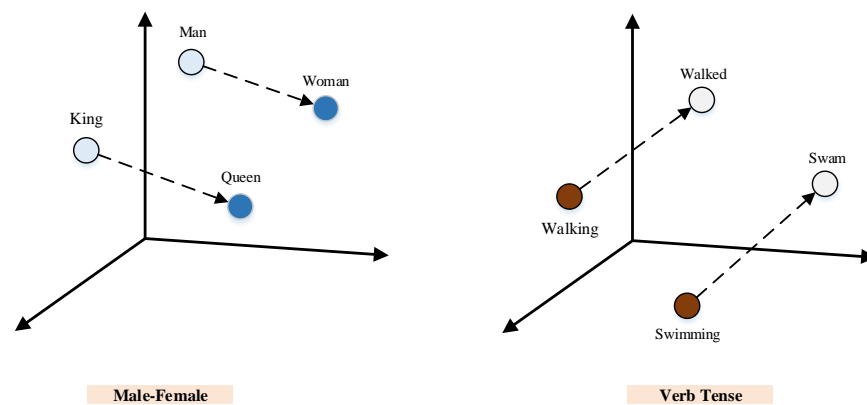


Figure 2. Word Embedding examples.

In this work, the Word2Vec word embedding model is performed. In Word2Vec, three layers are presented, an input layer, a hidden layer, and an output layer. However, this neural network architecture merely learns the weights of the hidden layer that are the embeddings of words.

Since Word2Vec requires huge datasets to operate, a pre-trained word embedding method is required to overcome the lack of such data. Therefore, we used a pre-trained word embedding corpus of 10 million instances [68]. It is worth mentioning that the number of dimensions for the learning data is 400; therefore, each collected dataset will have the same number. Five different datasets have been prepared, namely medical gloves, hand sanitizer, medical oxygen, and face masks, while the fifth dataset is composed of merging all four datasets. The details of the datasets can be found in Table 1.

Table 1. The description of the datasets.

Datasets	Tokenization	Product Type	Dimensions
Data 1	Word Embedding	Medical Gloves	400
Data 2	Word Embedding	Hand Sanitizer	400
Data 3	Word Embedding	Medical Oxygen	400
Data 4	Word Embedding	Face Masks	400
Data 5	Word Embedding	Merged all datasets	400

4.3. Proposed Approach

In this subsection, we explain the proposed approach based on harmony search algorithm and different classification models to obtain improved and competitive results using fewer features. The best subset of features is selected by using wrapper feature selection, where each classification model evaluates the output features.

In this study, the produced solution (feature subset) was designed as a set of binary vectors with length n , where n denotes the number of features in each dataset. The feature is considered selected if the corresponding value equals 1 and zero if not. The classification model evaluation (fitness) was used to determine the quality of a feature subset according to the following Equation (3):

$$Fitness = \alpha \times (1 - accuracy) + (1 - \alpha \times \frac{|S|}{|W|}) \quad (3)$$

where the number of selected features denotes by $|S|$, while $|W|$ presents the number of the total features in a given dataset, and $\alpha \in [0, 1]$.

Therefore, after completing the preparation of the data, we divided the data into training and testing sets. The splitting criteria used in this work was the 10-fold cross-

validation. To be more precise, the dataset is divided into nine training partitions, whereas the remaining partition is assigned for testing. Meanwhile, the number of features are listed in the HSA algorithm, thus identifying the total number of values in each solution. The HSA algorithm generates a solution and provides the classification models with a subset of features for the training phase.

Afterward, through an internal 5-fold cross-validation, the classifiers start training and building their models from the given training set. The purpose of performing this extra step is to improve training and build a more robust classifier model. To improve the next solution by HSA, the evaluation of the training set will be used as a fitness value in the HSA algorithm. A termination criterion is triggered when the HSA algorithm reaches its maximum number of iterations. In case of failure to meet this termination criterion, the fitness value will continue to be provided for the HSA algorithm. When the HSA reaches the maximum number of iterations, the best subset of features will be determined as the optimal solution. Based on the best subset of features obtained from the training phase, the final result of the classification accuracy and root mean square error (RMSE) for the testing set will be calculated using these subsets. Finally, all experiments were repeated 30 times independently to obtain reliable results, and the average was reported. The methodology is depicted in Figure 3 along with all the proposed steps.

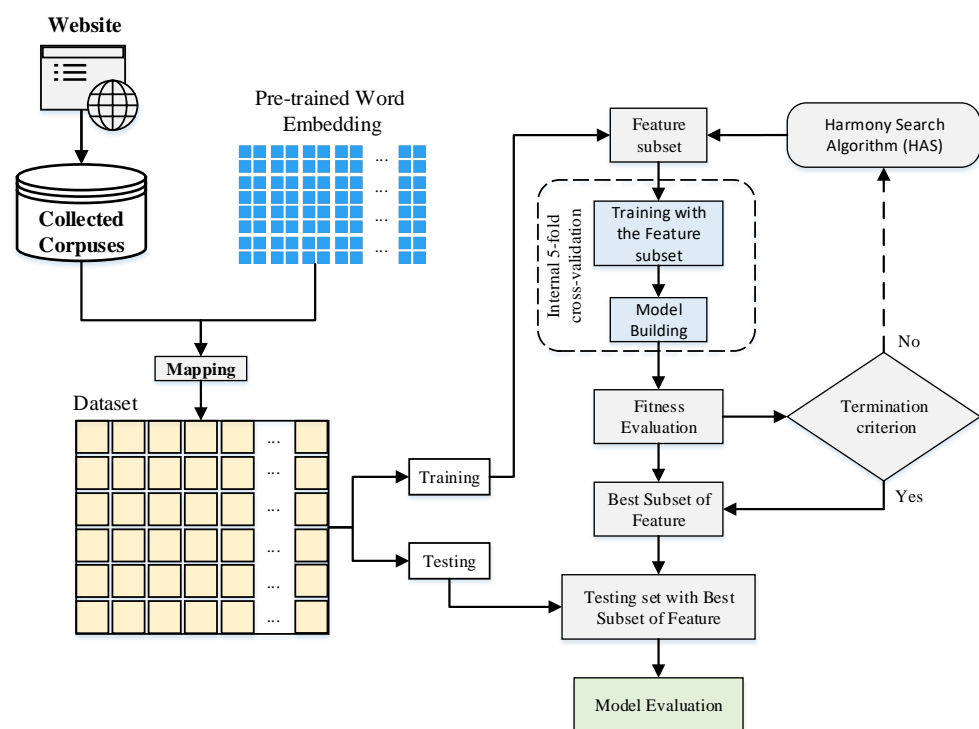


Figure 3. A detailed description of the methodology processes.

5. Experiment and Results

This section describes the findings from tests involving the HSA algorithm and other classification methods, both with and without feature selection. It also discusses the chosen assessment measures and the experiment's surrounding settings and conditions.

5.1. Experimental Setup

The experiments in this article were performed on a laptop running Windows 10 on an Intel Core i7 processor running at 2.40 GHz with 16 GB of RAM. Additionally, we ran our tests in the Matlab R2015a (8.5.0.197613) software environment. Further, the parameters of the classification models and HSA algorithm can be found in Table 2.

Table 2. Initial parameters of the classification models.

Algorithm	Parameter	Value
HSA	HMCR	[0, 1]
	PAR	[0, 1]
	bw	1.0
SVM	γ	[0.0001, 32.0]
	Cost (C)	[0.01, 35,000.0]
RF	Numbers of trees	100
AdaBoost	n_estimator	50
	learning_rate	1
Bagging	n_estimators	10
k-NN	k	1
Reptree	Tree size	100

5.2. Evaluation Measures

The evaluation measures considered in this study include the accuracy and RMSE for measuring the classification performance.

The accuracy reflects the quality of the classification task. It is defined as the ratio of the true positives (TP) and true negatives (TN) to the number of all predicted values including the TP, TN, false positives (FP), and true negatives (TN), as reflected by Equation (4) [69–74].

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

On the other hand, the RMSE measures the difference between each predicted output and the corresponding expected value by calculating the root of the ratio of the difference and the number of instances, as reflected by Equation (5).

$$RMSE = \sqrt{\frac{\sum(R_i - \bar{P}_i)^2}{n}}, \quad (5)$$

where n is the number of instances, R_i is the real value, P_i is the predicted value, and \bar{P}_i is the mean of the predicted values.

5.3. Comparison Experiments without Feature Selection

In this section, word embedding is experimented without any feature selection technique using the RF, K-NN, AdaBoost, bagging, SVM, and REPTree techniques. The experiments are made on different datasets, including the medical gloves, hand sanitizer, face masks, and whole datasets.

Table 3 shows the accuracy values achieved by conducting the results for the different datasets. In general, and according to the table, the Rf technique has the best results for the medical gloves, medical oxygen, and whole datasets. Looking deeply into the table, it shows that the RF technique has almost similar results to the bagging and REPTree techniques for all the datasets, and thus, the bagging and REPTree techniques slightly outperform the RF for the face masks and the hand sanitizer datasets, respectively. On the other hand, the AdaBoost technique also has similar results to RF for the hand sanitizer and medical oxygen datasets. Moreover, K-NN has a close accuracy value to RF for the whole datasets but is not competitive to Rf for the remaining datasets. SVM, on the other hand, has a non-competitive behavior to RF and almost all the other techniques. It also has a very low accuracy value for the face masks compared to the others. In summary, RF, REPTree, and bagging tend to have consistent behavior across all datasets, having better results than the other techniques. However, RF is recommended, since it has better results than bagging and REPTree for the majority of the datasets, although it shows slight enhancement.

Table 4 shows the results obtained by RMSE for the same datasets and techniques. Similar observations can be made for RF, bagging, and REPtree techniques. However, k-NN shows a slight enhancement for the face masks dataset. SVM also shows the worst results among all the datasets.

Table 3. Accuracy results of all models for all datasets with word embedding technique. Bold font indicates best result.

Data	RF	k-NN	AdaBoost	Bagging	SVM	REPtree
Medical Gloves	87.489	84.412	79.044	87.364	79.044	86.074
Hand Sanitizer	84.034	76.244	83.340	83.790	80.039	84.261
Medical Oxygen	81.395	73.924	81.383	81.182	79.373	80.552
Face Masks	90.821	88.639	73.086	90.903	59.500	90.650
Whole Datasets	93.458	92.032	77.392	93.050	75.822	93.431

Table 4. RMSE results of all models for all datasets with word embedding technique. Bold font indicates best result.

Data	RF	k-NN	AdaBoost	Bagging	SVM	REPtree
Medical Gloves	0.230	0.251	0.312	0.231	0.323	0.231
Hand Sanitizer	0.235	0.273	0.273	0.237	0.283	0.237
Medical Oxygen	0.249	0.281	0.287	0.250	0.287	0.254
Face Masks	0.238	0.237	0.392	0.240	0.520	0.224
Whole Datasets	0.155	0.160	0.309	0.157	0.311	0.148

5.4. Evolutionary Algorithm Feature Selection

This section discusses the effect of adding the feature selection process using the HSA technique for different classification techniques including RF, k-NN, AdaBoost, bagging, SVM, and REPtree for the five datasets in the study.

Table 5 shows the effect of the HSA on reducing the dimensions of the datasets. We can observe that the medical glove, hand sanitizer, medical oxygen, face masks, and whole datasets have reduced dimensions of 23, 42, 43, 30, and 63, respectively. By comparing the values to the ones in Table 1, where every dataset has a dimension value of 400, it is observed that the dimension value has been highly reduced for all the datasets. Specifically, the number of features is reduced by 94.25%, 89.5%, 89.25%, 92.5%, and 84.25% for the medical glove, hand sanitizer, medical oxygen, face masks, and whole datasets, respectively.

Table 5. The number of reduced dimensions for each dataset using the HSA technique.

Datasets	Reduced Dimensions
Medical Gloves	23
Hand Sanitizer	42
Medical Oxygen	43
Face Masks	30
Whole Datasets	63

According to the reduced dimensions of the different datasets, Tables 6 and 7 show the obtained results of the accuracy and RMSE values, respectively, from experimenting with the same aforementioned classification techniques.

On one hand, Table 6 shows the accuracy values for all the datasets. It can also be concluded that the RF technique performed better than the other techniques. The

same observation can be made for the bagging and the REPTree techniques, having close results to the RF technique. However, different observations can be made for the k-NN technique. This includes the enhanced accuracy values for the k-NN in all the datasets compared to the performance achieved by k-NN without considering the feature selection process, which can be recalled from Table 3. It also includes an identical performance for the K-NN compared to the RF for the medical oxygen dataset and almost similar results compared to RF for all the datasets. In addition, AdaBoost made the same observation for the hand sanitizer and medical oxygen datasets, having very close results to those for the RF technique. SVM still achieves the worst results among all other techniques.

On the other hand, Table 7 shows the RMSE values for all the datasets. Different observations can be made from this table, by the enhanced results observed by k-NN, having better performance than the other techniques, including RF, for medical gloves, medical oxygen, and face masks datasets. Bagging slightly achieved better results for the hand sanitizer dataset. In general, REPTree, and k-NN have similar results, as observed from Table 6. In contrast, SVM shows the worst results among the other techniques and a relatively very high RMSE value for the face masks dataset.

To summarize, the process of reducing the number of features for the dataset does not only allow for low processing time achieved by the classification techniques but also enhances the performance of some low-performance techniques including the k-NN and keeping a competitive performance for the other classification techniques.

Table 6. Accuracy results of all models for all datasets with word embedding and feature selection techniques. Bold font indicates best result.

Data	HSA-RF	HSA-k-NN	HSA-AdaBoost	HSA-Bagging	HSA-SVM	HSA-REPTree
Medical Gloves	87.022	86.397	79.044	86.397	79.044	86.074
Hand Sanitizer	84.546	84.032	83.333	84.371	80.039	84.487
Medical Oxygen	81.255	81.255	81.381	80.877	79.373	80.716
Face Masks	90.389	90.381	73.127	91.508	59.500	90.650
Whole Datasets	93.691	93.215	77.996	93.426	75.822	93.458

Table 7. RMSE results of all models for all datasets with word embedding and feature selection techniques. Bold font indicates best result.

Data	HSA-RF	HSA-k-NN	HSA-AdaBoost	HSA-Bagging	HSA-SVM	HSA-REPTree
Medical Gloves	0.231	0.228	0.311	0.235	0.323	0.231
Hand Sanitizer	0.234	0.236	0.273	0.233	0.283	0.235
Medical Oxygen	0.251	0.251	0.286	0.252	0.287	0.252
Face Masks	0.238	0.221	0.392	0.240	0.520	0.224
Whole Datasets	0.154	0.153	0.309	0.156	0.311	0.148

5.5. Discussion

The proposed approach consists of using word embedding as text representation and a harmony search algorithm as a feature selection method utilizing different classification models as an evaluation to improve the sentiment analysis prediction of medical products.

The implementation of word embedding as a text representation can benefit by selecting words with symmetrical meanings to have similar representations. Thus, reducing the number of presenting words such as the methods in Figure 4. We can notice from the table that some datasets reached nearly 7500 features, where most of these features are useless, unlike the word embedding method. Therefore, using word embedding can

be more efficient, especially using the pre-trained word embedding technique through mapping our data with a learning corpus.

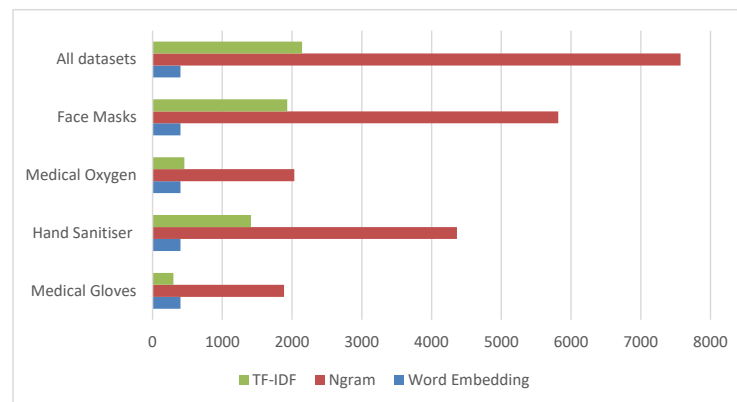


Figure 4. Feature number of other text representation methods.

Moreover, the results of the proposed approach in the previous subsection highlight the improvement of using the HSA for feature selection. Such a mechanism showed excellent and competitive results using a fewer number of features, consequently decreasing the computational time of the experimental results. The reduction rates on the features are shown in Figure 5. It can be seen that the percentage of reduction reached about 90% on every dataset.

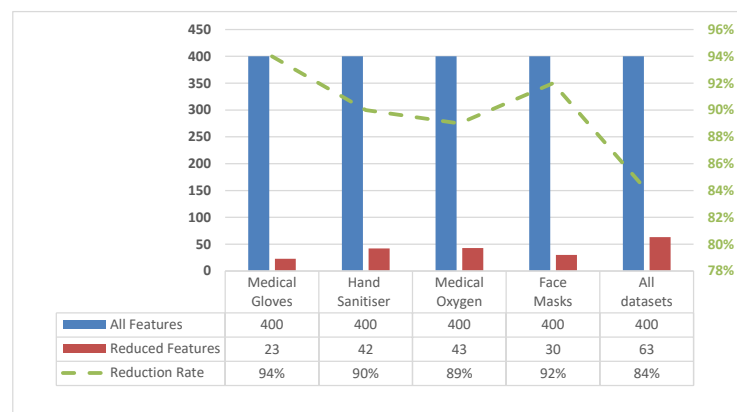


Figure 5. Features reduction rates.

6. Conclusions

This study proposes an intelligent hybrid technique using a harmony search algorithm for feature selection and different classifiers for evaluation. This paper studies customers' reviews regarding COVID-19 medical products on the Alibaba e-commerce website. Five different datasets have been prepared, namely medical gloves, hand sanitizer, medical oxygen, and face masks, while the fifth dataset is composed of merging all four datasets. After data collection and preprocessing, an advance pre-trained word embedding technique was applied for the tokenization process; the Word2Vec embedding technique was chosen as the main feature extraction model rather than the regular techniques to overcome the challenges of the data. This work follows a hybrid evolutionary approach, as it is based on a metaheuristic optimization algorithm, namely a harmony search algorithm, which is used for the feature selection process to select the best subset of features using different classifiers for evaluation. Six different classification techniques including RF, k-NN, AdaBoost, bagging, SVM, and REPTree are examined for classifying the sentiment dataset. The results include the accuracy and RMSE values of the classification process with the feature selection process and without it. The results show that applying the feature

selection process reduced the number of features by 94.25%, 89.5%, 89.25%, 92.5%, and 84.25% for the medical glove, hand sanitizer, medical oxygen, face masks, and whole datasets, respectively. Competitive performance for the classification techniques and a decrease of the computational time of the classification process due to reducing the number of features are achieved by improving the evaluation measures' values of some low-performance techniques.

Author Contributions: Conceptualization, R.O. and A.M.A.-Z.; Methodology, R.O., A.M.A.-Z.; Validation, R.O. and A.M.A.-Z.; Data curation, R.O. and A.M.A.-Z.; Writing original draft preparation, R.O., O.H., R.Q., L.A.-Q. and A.M.A.-Z.; Writing review and editing, R.O., O.H., R.Q., L.A.-Q. and A.M.A.-Z.; Supervision R.O., O.H. and A.M.A.-Z.; Project administration, R.O. and A.M.A.-Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Aljarah, I.; Habib, M.; Hijazi, N.; Faris, H.; Qaddoura, R.; Hammo, B.; Abushariah, M.; Alfawareh, M. Intelligent detection of hate speech in Arabic social network: A machine learning approach. *J. Inf. Sci.* **2020**, *47*, 0165551520917651. [\[CrossRef\]](#)
- Al-Zoubi, A.; Alqatawna, J.; Faris, H.; Hassonah, M.A. Spam profiles detection on social networks using computational intelligence methods: The effect of the lingual context. *J. Inf. Sci.* **2021**, *47*, 58–81. [\[CrossRef\]](#)
- Injadat, M.; Salo, F.; Nassif, A.B. Data mining techniques in social media: A survey. *Neurocomputing* **2016**, *214*, 654–670. [\[CrossRef\]](#)
- Jiménez-Zafra, S.M.; Martín-Valdivia, M.T.; Molina-González, M.D.; Ureña-López, L.A. How do we talk about doctors and drugs? Sentiment analysis in forums expressing opinions for medical domain. *Artif. Intell. Med.* **2019**, *93*, 50–57. [\[CrossRef\]](#) [\[PubMed\]](#)
- Faris, H.; Alqatawna, J.; Ala'M, A.Z.; Aljarah, I. Improving email spam detection using content based feature engineering approach. In Proceedings of the 2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), Aqaba, Jordan, 11–13 October 2017; pp. 1–6.
- Medhat, W.; Hassan, A.; Korashy, H. Sentiment analysis algorithms and applications: A survey. *Ain Shams Eng. J.* **2014**, *5*, 1093–1113. [\[CrossRef\]](#)
- Beigi, G.; Hu, X.; Maciejewski, R.; Liu, H. An overview of sentiment analysis in social media and its applications in disaster relief. *Sentim. Anal. Ontol. Eng.* **2016**, *639*, 313–340.
- Harfoushi, O.; Hasan, D.; Obiedat, R. Sentiment analysis algorithms through azure machine learning: Analysis and comparison. *Mod. Appl. Sci.* **2018**, *12*, 49. [\[CrossRef\]](#)
- Altawaier, M.M.; Tiun, S. Comparison of machine learning approaches on arabic twitter sentiment analysis. *Int. J. Adv. Sci. Eng. Inf. Technol.* **2016**, *6*, 1067–1073. [\[CrossRef\]](#)
- Gautam, G.; Yadav, D. Sentiment analysis of twitter data using machine learning approaches and semantic analysis. In Proceedings of the 2014 Seventh International Conference on Contemporary Computing (IC3), Noida, India, 7–9 August 2014; pp. 437–442.
- Samal, B.; Behera, A.K.; Panda, M. Performance analysis of supervised machine learning techniques for sentiment analysis. In Proceedings of the 2017 Third International Conference on Sensing, Signal Processing and Security (ICSSS), Chennai, India, 4–5 May 2017; pp. 128–133.
- Ahelegbey, D.F.; Cerchiello, P.; Scaramozzino, R. Network Based Evidence of the Financial Impact of Covid-19 Pandemic. 2021. Available online: <https://ideas.repec.org/p/pav/demwpp/demwp0198.html> (accessed on 1 November 2021).
- Gonçalves, P.; Araújo, M.; Benevenuto, F.; Cha, M. Comparing and combining sentiment analysis methods. In Proceedings of the First ACM Conference on Online Social Networks, Boston, MA, USA, 7–8 October 2013; pp. 27–38.
- Feldman, R. Techniques and applications for sentiment analysis. *Commun. ACM* **2013**, *56*, 82–89. [\[CrossRef\]](#)
- Hasan, A.; Moin, S.; Karim, A.; Shamshirband, S. Machine learning-based sentiment analysis for twitter accounts. *Math. Comput. Appl.* **2018**, *23*, 11. [\[CrossRef\]](#)
- Oueslati, O.; Khalil, A.I.S.; Ounelli, H. Sentiment analysis for helpful reviews prediction. *Int. J.* **2018**, *7*, 34–40.
- Sharef, N.M.; Zin, H.M.; Nadali, S. Overview and Future Opportunities of Sentiment Analysis Approaches for Big Data. *J. Comput. Sci.* **2016**, *12*, 153–168. [\[CrossRef\]](#)
- Denecke, K.; Deng, Y. Sentiment analysis in medical settings: New opportunities and challenges. *Artif. Intell. Med.* **2015**, *64*, 17–27. [\[CrossRef\]](#) [\[PubMed\]](#)
- Yadav, S.; Ekbal, A.; Saha, S.; Bhattacharyya, P. Medical sentiment analysis using social media: Towards building a patient assisted system. In Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), Miyazaki, Japan, 7–12 May 2018.
- Rozenblum, R.; Bates, D.W. Patient-Centred Healthcare, Social Media and the Internet: The Perfect Storm? 2013. Available online: <https://qualitysafety.bmj.com/content/22/3/183> (accessed on 1 November 2021).

21. Abualigah, L.; Alfar, H.E.; Shehab, M.; Hussein, A.M.A. Sentiment analysis in healthcare: A brief review. In *Recent Advances in NLP: The Case of Arabic Language*; Springer: Cham, Switzerland, 2020; pp. 129–141.
22. Polisen, J.; Andellini, M.; Salerno, P.; Borsci, S.; Pecchia, L.; Iadanza, E. Case Studies on the Use of Sentiment Analysis to Assess the Effectiveness and Safety of Health Technologies: A Scoping Review. *IEEE Access* **2021**, *9*, 66043–66051. [[CrossRef](#)]
23. Li, J.; Xu, Q.; Cuomo, R.; Purushothaman, V.; Mackey, T. Data mining and content analysis of the Chinese social media platform Weibo during the early COVID-19 outbreak: Retrospective observational infoveillance study. *JMIR Public Health Surveill.* **2020**, *6*, e18700. [[CrossRef](#)]
24. Espitia, A.; Rocha, N.; Ruta, M. Trade in Critical Covid-19 Products. 2020. Available online: <https://openknowledge.worldbank.org/handle/10986/33514> (accessed on 1 November 2021).
25. Kampf, G.; Scheithauer, S.; Lemmen, S.; Saliou, P.; Suchomel, M. COVID-19-associated shortage of alcohol-based hand rubs, face masks, medical gloves, and gowns: Proposal for a risk-adapted approach to ensure patient and healthcare worker safety. *J. Hosp. Infect.* **2020**, *105*, 424–427. [[CrossRef](#)]
26. Deshpande, M.; Sarkar, A. BI and sentiment analysis. *Bus. Intell. J.* **2010**, *15*, 41–49.
27. Gereffi, G. What does the COVID-19 pandemic teach us about global value chains? The case of medical supplies. *J. Int. Bus. Policy* **2020**, *3*, 287–301. [[CrossRef](#)]
28. Ab Rahman, M.F.; Rusli, A.; Misman, M.A.; Rashid, A.A. Biodegradable gloves for waste management post-COVID-19 outbreak: A shelf-life prediction. *ACS Omega* **2020**, *5*, 30329–30335. [[CrossRef](#)]
29. Stein, F.; Perry, M.; Banda, G.; Woolhouse, M.; Mutapi, F. Oxygen provision to fight COVID-19 in sub-Saharan Africa. *BMJ Glob. Health* **2020**, *5*, e002786. [[CrossRef](#)]
30. Nakkazi, E. Oxygen supplies and COVID-19 mortality in Africa. *Lancet Respir. Med.* **2021**, *9*, e39. [[CrossRef](#)]
31. Sardesai, I.; Grover, J.; Garg, M.; Nanayakkara, P.; Di Somma, S.; Paladino, L.; Anderson, H.L., III; Gaieski, D.; Galwankar, S.C.; Stawicki, S.P. Short term home oxygen therapy for COVID-19 patients: The COVID-HOT algorithm. *J. Fam. Med. Prim. Care* **2020**, *9*, 3209. [[CrossRef](#)]
32. Hajiali, M. Big data and sentiment analysis: A comprehensive and systematic literature review. *Concurr. Comput. Pract. Exp.* **2020**, *32*, e5671. [[CrossRef](#)]
33. Sun, Y.; Bao, Q.; Lu, Z. Coronavirus (Covid-19) outbreak, investor sentiment, and medical portfolio: Evidence from China, Hong Kong, Korea, Japan, and US. *Pac.-Basin Financ. J.* **2021**, *65*, 101463. [[CrossRef](#)]
34. Yi, S.; Liu, X. Machine learning based customer sentiment analysis for recommending shoppers, shops based on customers' review. *Complex Intell. Syst.* **2020**, *6*, 621–634. [[CrossRef](#)]
35. Obiedat, R.; Harfoushi, O.; Qaddoura, R.; Al-Qaisi, L.; Al-Zoubi, A.M. An Evolutionary-Based Sentiment Analysis Approach for Enhancing Government Decisions during COVID-19 Pandemic: The Case of Jordan. *Appl. Sci.* **2021**, *11*, 9080. [[CrossRef](#)]
36. Kumar, A.; Jaiswal, A. Systematic literature review of sentiment analysis on Twitter using soft computing techniques. *Concurr. Comput. Pract. Exp.* **2020**, *32*, e5107. [[CrossRef](#)]
37. Jain, P.K.; Pamula, R.; Srivastava, G. A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews. *Comput. Sci. Rev.* **2021**, *41*, 100413. [[CrossRef](#)]
38. Chaturvedi, I.; Cambria, E.; Welsch, R.E.; Herrera, F. Distinguishing between facts and opinions for sentiment analysis: Survey and challenges. *Inf. Fusion* **2018**, *44*, 65–77. [[CrossRef](#)]
39. Gohil, S.; Vuik, S.; Darzi, A. Sentiment analysis of health care tweets: Review of the methods used. *JMIR Public Health Surveill.* **2018**, *4*, e5789. [[CrossRef](#)]
40. Lagu, T.; Metayer, K.; Moran, M.; Ortiz, L.; Priya, A.; Goff, S.L.; Lindenauer, P.K. Website characteristics and physician reviews on commercial physician-rating websites. *JAMA* **2017**, *317*, 766–768. [[CrossRef](#)]
41. Liu, J.; Zhou, Y.; Jiang, X.; Zhang, W. Consumers' satisfaction factors mining and sentiment analysis of B2C online pharmacy reviews. *BMC Med. Inform. Decis. Mak.* **2020**, *20*, 1–13. [[CrossRef](#)] [[PubMed](#)]
42. Na, J.C.; Kyaing, W.Y.M. Sentiment analysis of user-generated content on drug review websites. *J. Inf. Sci. Theory Pract.* **2015**, *3*, 6–23. [[CrossRef](#)]
43. Gräßler, F.; Kallumadi, S.; Malberg, H.; Zaunseder, S. Aspect-based sentiment analysis of drug reviews applying cross-domain and cross-data learning. In Proceedings of the 2018 International Conference on Digital Health, Lyon, France, 23–26 April 2018; pp. 121–125.
44. Daniulaityte, R.; Chen, L.; Lamy, F.R.; Carlson, R.G.; Thirunarayan, K.; Sheth, A. "When 'bad' is 'good'": Identifying personal communication and sentiment in drug-related tweets. *JMIR Public Health Surveill.* **2016**, *2*, e6327. [[CrossRef](#)] [[PubMed](#)]
45. Basiri, M.E.; Abdar, M.; Cifci, M.A.; Nemati, S.; Acharya, U.R. A novel method for sentiment classification of drug reviews using fusion of deep and machine learning techniques. *Knowl.-Based Syst.* **2020**, *198*, 105949. [[CrossRef](#)]
46. Harrison, C.J.; Sidey-Gibbons, C.J. Machine learning in medicine: A practical introduction to natural language processing. *BMC Med. Res. Methodol.* **2021**, *21*, 1–11. [[CrossRef](#)]
47. Karyotaki, M.; Drigas, A.; Skianis, C. Attentional control and other executive functions. *Int. J. Emerg. Technol. Learn.* **2017**, *12*, 219–233.
48. Drigas, A.; Karyotaki, M.; Skianis, C. Success: A 9 Layered-based Model of Giftedness. *Int. J. Recent Contrib. Eng. Sci. IT* **2017**, *5*, 4–18. [[CrossRef](#)]

49. Papoutsis, C.; Drigas, A.; Skianis, C. Emotional intelligence as an important asset for HR in organizations: Attitudes and working variables. *Int. J. Adv. Corp. Learn.* **2019**, *12*, 21. [[CrossRef](#)]
50. Bürkner, P.C.; Vuorre, M. Ordinal Regression Models in Psychology: A Tutorial. *Adv. Methods Pract. Psychol. Sci.* **2019**, *2*, 77–101. [[CrossRef](#)]
51. Gutiérrez, P.A.; Pérez-Ortiz, M.; Sánchez-Monedero, J.; Fernández-Navarro, F.; Hervás-Martínez, C. Ordinal Regression Methods: Survey and Experimental Study. *IEEE Trans. Knowl. Data Eng.* **2016**, *28*, 127–146. [[CrossRef](#)]
52. Eiben, A.E.; Smith, J.E. What is an evolutionary algorithm? In *Introduction to Evolutionary Computing*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 25–48.
53. Qaddoura, R.; Faris, H.; Aljarah, I. An efficient evolutionary algorithm with a nearest neighbor search technique for clustering analysis. *J. Ambient Intell. Humaniz. Comput.* **2021**, *12*, 8387–8412. [[CrossRef](#)]
54. Sloss, A.N.; Gustafson, S. 2019 Evolutionary Algorithms Review. *arXiv* **2019**, arXiv:1906.08870.
55. Qaddoura, R.; Manaseer, W.A.; Abushariah, M.A.; Alshraideh, M.A. Dental radiography segmentation using expectation-maximization clustering and grasshopper optimizer. *Multimed. Tools Appl.* **2020**, *79*, 22027–22045. [[CrossRef](#)]
56. Abualigah, L.; Diabat, A.; Geem, Z.W. A comprehensive survey of the harmony search algorithm in clustering applications. *Appl. Sci.* **2020**, *10*, 3827. [[CrossRef](#)]
57. Al-Omouh, A.A.; Alsewari, A.A.; Alamri, H.S.; Zamli, K.Z. Comprehensive Review of the Development of the Harmony Search Algorithm and its Applications. *IEEE Access* **2019**, *7*, 14233–14245. [[CrossRef](#)]
58. Ala'M, A.Z.; Hassonah, M.A.; Heidari, A.A.; Faris, H.; Mafarja, M.; Aljarah, I. Evolutionary competitive swarm exploring optimal support vector machines and feature weighting. *Soft Comput.* **2021**, *25*, 3335–3352.
59. Gokalp, O.; Tasci, E.; Ugur, A. A novel wrapper feature selection algorithm based on iterated greedy metaheuristic for sentiment classification. *Expert Syst. Appl.* **2020**, *146*, 113176. [[CrossRef](#)]
60. Ahmed, S.; Mafarja, M.; Faris, H.; Aljarah, I. Feature selection using salp swarm algorithm with chaos. In Proceedings of the 2nd International Conference on Intelligent Systems, Metaheuristics & Swarm Intelligence, Phuket, Thailand, 24–25 March 2018; pp. 65–69.
61. Kumar, H.; Harish, B.; Darshan, H. Sentiment Analysis on IMDb Movie Reviews Using Hybrid Feature Extraction Method. *Int. J. Interact. Multimed. Artif. Intell.* **2019**, *5*. [[CrossRef](#)]
62. Salau, A.O.; Jain, S. Feature Extraction: A Survey of the Types, Techniques, Applications. In Proceedings of the 2019 International Conference on Signal Processing and Communication (ICSC), Noida, India, 7–9 March 2019; pp. 158–164. [[CrossRef](#)]
63. Li, Y.; Yang, T. Word embedding for understanding natural language: A survey. In *Guide to Big Data Applications*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 83–104.
64. Srinivasan, S.; Ravi, V.; Alazab, M.; Ketha, S.; Ala'M, A.Z.; Padannayil, S.K. Spam emails detection based on distributed word embedding with deep learning. In *Machine Intelligence and Big Data Analytics for Cybersecurity Applications*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 161–189.
65. Kaibi, I.; Nfaoui, E.H.; Satori, H. Sentiment analysis approach based on combination of word embedding techniques. In *Embedded Systems and Artificial Intelligence*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 805–813.
66. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient estimation of word representations in vector space. *arXiv* **2013**, arXiv:1301.3781.
67. Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G.S.; Dean, J. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems*; Curran Associates Inc.: Lake Tahoe, NV, USA, 2013; pp. 3111–3119.
68. Petrović, S.; Osborne, M.; Lavrenko, V. The edinburgh twitter corpus. In Proceedings of the NAACL HLT 2010 Workshop on Computational Linguistics in a World of Social Media, Los Angeles, CA, USA, 6 June 2010; pp. 25–26.
69. Qaddoura, R.; Aljarah, I.; Faris, H.; Almomani, I. A Classification Approach Based on Evolutionary Clustering and Its Application for Ransomware Detection. In *Evolutionary Data Clustering: Algorithms and Applications*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 237–248.
70. Shaw, S.; Prakash, M. Solar Radiation Forecasting Using Support Vector Regression. In Proceedings of the 2019 International Conference on Advances in Computing and Communication Engineering (ICACCE), Sathyamangalam, India, 4–6 April 2019; pp. 1–4.
71. Ala'M, A.Z.; Rodan, A.; Alazzam, A. Classification model for credit data. In Proceedings of the 2018 Fifth HCT Information Technology Trends (ITT), Dubai, United Arab Emirates, 28–29 November 2018; pp. 132–137.
72. Yaghi, R.I.; Faris, H.; Aljarah, I.; Ala'M, A.Z.; Heidari, A.A.; Mirjalili, S. Link prediction using evolutionary neural network models. In *Evolutionary Machine Learning Techniques*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 85–111.
73. Qaddoura, R.; Al-Zoubi, A.; Almomani, I.; Faris, H. A Multi-Stage Classification Approach for IoT Intrusion Detection Based on Clustering with Oversampling. *Appl. Sci.* **2021**, *11*, 3022. [[CrossRef](#)]
74. Qaddoura, R.; Al-Zoubi, M.; Faris, H.; Almomani, I. A Multi-Layer Classification Approach for Intrusion Detection in IoT Networks Based on Deep Learning. *Sensors* **2021**, *21*, 2987. [[CrossRef](#)] [[PubMed](#)]