

Classification of Twitter Data Belonging to Sudanese Revolution Using Text Mining Techniques

Huda Jamal Abdelhameed, Susana Muñoz- Hernández

Abstract- Twitter has attracted a great deal of attention recently. It is one of the most common social networking sites for chats, sharing ideas, and transfer of information and news through text. This paper focused on sentiment classification of twitter data belonging to Sudanese revolution which written using either Modern Standard Arabic or Sudanese dialectal Arabic. Twitter's API was used to collect tweets related to Sudanese revolution. The dataset consists of 6482 tweets with a good balance of positive and negative sentiments. Three different classifiers were used on the dataset namely; Support Vector Machine (SVM), Naive Bayes (NB) and Decision Tree (DT) to classify the tweets based on its polarity into positive or negative. We evaluated our work by four different measures which are Precision, Recall, Accuracy and F-measure. We made a comparison between the three classifiers based on those measures. The results show that, SVM achieved the best Accuracy and F-measure and it equals 75.2%, 83.9% respectively. While NB achieved best Precision and it equals 75.2%. Also, DT achieved best Recall and it equals 99.9%. In addition, the percentages of positive and negative opinions toward the government was calculated. 9.4% represents the percentage of positive opinions related the government, while 90.6% represents the percentage of negative opinions related the same government.

Index Terms— Sentiment Classification; Opinion Mining; Sentiment Analysis; Twitter Data; Sudanese Revolution; Sudanese Dialect.

I. INTRODUCTION

Opinion mining or sentiment analysis is a technique to detect and extract subjective information in text documents. In general, sentiment analysis tries to determine the sentiment of a writer about some aspect or the overall contextual polarity of a document. The sentiment may be his or her judgment, mood or evaluation [1]. Sentiment Analysis focus on study people's opinions, attitudes and emotions. In order to promote marketing, large companies and business, people are making use of opinion mining [2]. Sentiment analysis is rapidly growing area. On the other hand, the users of social media post a huge amount of comments on a daily basis that's why an [analysis of social media has attracted a great deal of attention recently. There are several challenges in Sentiment analysis. The first is an opinion word that is considered to be positive in one situation may be considered negative in another situation. A second challenge is that people don't

always express opinions in a same way. Most traditional text processing relies on the fact that small differences between two pieces of text don't change the meaning very much. In Sentiment analysis, however, "the picture was great" is very different from "the picture was not great". People can be contradictory in their statements. Most reviews will have both positive and negative comments, which is somewhat manageable by analyzing sentences one at a time [3]. Sentiment classification is a key problem in sentiment analysis, where a text is labeled as a positive or negative evaluation of a particular object.

There are two main classification techniques, supervised and unsupervised. Supervised classification techniques are also known as predictive or directed classification. In this method set of possible class is known in advanced. Unsupervised classification techniques are also known as descriptive or undirected. In this method set of possible class is unknown, after classification we can assign name to that class [4]. Before applying any algorithm, pre-processing step is carried out, such as tokenization, stemming, normalization, stopword removal. Opinion mining has three levels of granularities: Document level, Sentence level and Aspect level. In this work we focused on sentence level of sentiment classification.

II. RELATED WORKS

(A) Hissah AL-Saif, Hmood Al-Dossari [4] have used classification techniques to detect crimes and identify their nature of different classification algorithms. Their experiments evaluated different algorithms, such as SVM, DT, CNB, and KNN, in terms of accuracy and speed in the crime domain. Also, different feature extraction techniques are evaluated, including root-based stemming, light stemming, n-gram. Their experiments revealed the superiority of n-gram over other techniques. Specifically, the results indicate the superiority of SVM with tri-gram over other classifiers, with a 91.55% accuracy.

(B) Achin Jain, Vanita Jain [5] discussed about the sentimental analysis process. They carried out sentiment analysis and classification task of tweets belonging to #Renewable Energy. They applied five different machine learning algorithms for the classification of tweets into three categories. They have carried classification without feature selection technique and with feature selection techniques. They have used CfsSubsetEvaluation and Information Gain feature selection methods to reduce the number of features from the dataset. Their results show that the accuracy of sentiment classification is better with feature selection

Huda Jamal Abdelhameed, Department of Computer Systems and Networks, College of Computer Science and Information Technology, Sudan University of Science and Technology, Khartoum, Sudan.

Susana Munoz – Hernández, Escuela Técnica Superior de Ingenieros Informáticos Campus de Montegancedo, Boadilla del Monte, Madrid, Spain

methods. The best accuracy (92.96%) is achieved with Support Vector Machine (Using PUK Kernel) and CfsSubsetEval feature selection method.

(C) Ali Mustafa Qamar et al. [6] Presented sentiment analysis of tweets written in English, belonging to different telecommunication companies in Saudi Arabia. They apply different machine learning algorithms such as a k nearest neighbor algorithm, Artificial Neural Networks (ANN), Naïve Bayesian etc. They classified the tweets into positive, negative and neutral classes based on Euclidean distance as well as cosine similarity. Moreover, they also learned similarity matrices for KNN classification. fsSubsetEvaluation as well as Information Gain was used for feature selection. Their results of CfsSubsetEvaluation were better than the ones obtained with Information Gain. Moreover, in their study, KNN performed better than the other algorithms and gave 75.4%, 76.6% and 75.6% for Precision, Recall and F-measure, respectively. Furthermore, interesting trends wrt days, months etc. Was also discovered.

(D) Tagwa M. [7] presented some of the previous works in sentiment analysis by using two techniques: a lexicon-based technique and a Corpus-based technique. They addressed some experiments and studies that deal with sentiment analysis in Arabic. Their study aims to use sentiment classification for Arabic tweets around Khartoum. They used different techniques for Arabic sentiment analysis applied in Arabic tweets around Khartoum and decide if the sentiment is happiness (positive), sadness (negative) or neutral. They were created a corpus of Arabic tweets around Khartoum. Then build a lexicon for Arabic words. This lexicon contains a total of words divided in two groups; the words indicating happiness (positive) and sadness words (negative) with experts in language. They used two types of classification techniques, SVM and naive Bayes.

(E) Afnan A. Al-Subaihin et al. [8] proposed a lexicon-based sentiment analysis tool for colloquial Arabic text used in chatting, daily conversation and within social media. They have an independent component in their work which is game-based lexicons, that are based on human expertise. SVM, Naïve Bayes and Maximum Entropy classifiers are used in this study. However, they have proven that, SVM achieved the higher accuracy. Their tool should rely partially texts based on human judgment to overcome the problem arise from using non-standardized co colloquial Arabic text.

(F) Mohammed N. Al-Kabi et al. [9] proposed an opinion tool for analysis Most Standard Arabic and colloquial Arabic comments. Their tool is capable to identify the polarity, subjectivity, and strength of each comment. They build 18 lexicons manually. Two general purpose lexicons were built to identify polarity, and 16 domain-specific lexicons were built to identify the polarity with eight different domains: Technology, Books, Education, Movies, Places, Politics, Products and Society. They used Naive Bayes to classify the domain of the comments. Their experiments showed that the proposed tool yields more accurate results when it is applied on domain-based Arabic comments relative to general- based Arabic comments. As they present the tool yield 93.9% accuracy to classify the comments into their proper domains, a 90% accuracy to identify the real polarity, and a 96.9% accuracy to identify the strength of the comments. This study

used a small dataset, and the proposed tool is incapable to deal with emoticons and chat language.

III. TOOLS AND CLASSIFICATION TECHNIQUES

In this work RapidMiner used as a tool and three different classification techniques were applied which are Support Vector Machine (SVM), Naive Bayes (NB) and Decision Tree (DT).

A) RapidMiner

RapidMiner is a software platform developed by the company of the same name that provides an integrated environment for machine learning, data mining, text mining, predictive analytics and business analytics. It is used for business and commercial applications as well as for research, education, training, rapid prototyping, and application development and supports all steps of the data mining process including data preparation, results visualization, validation and optimization. RapidMiner is developed on an open core model, with the RapidMiner Basic Edition available for download under the AGPL license [10]. RapidMiner includes many operators that support text mining such as Text Processing package.

B) Classification Techniques

In this work we have used three different type of classification techniques which are Support Vector Machine (SVM), Naive Bayes (NB) and Decision Tree (DT).

1. Support Vector Machines classification approach:

Support Vector Machines (SVM) is one of the discriminative classification approaches which is commonly recognized to be more accurate. SVM classification approach is based on Structural Risk Minimization (SRM) principle from statistical learning theory. SRM is an inductive principle for model selection used for learning from finite training data and it provides a method for controlling the generalization ability of learning machines that uses a small size training data. The idea of this principle is to find a hypothesis to guarantee the lowest true error. In addition to this, the derivation of SVM is mathematically rigorous and very open to theoretical understanding and analysis.

SVM needs both positive and negative training datasets which are uncommon for other classification methods. It is outstanding from the others with its better classification performance and its ability in handling documents with high-dimensional input space and culls out most of the irrelevant features. The good generalization characteristic of SVM is due to the implementation of SRM which entails finding an optimal hyper-plane, thus guaranteeing the lowest classification error. Besides, a capacity which is independent of the dimensionality of the feature space makes SVM a highly accurate classifier in most applications. However, the major drawback of SVM is its relatively complex training and categorizing algorithms and also the high time and memory consumptions during the training stage and classifying stage due to its convoluted training and categorizing algorithms. Besides, confusions occur during the classification tasks because the documents could be annotated to several categories because of similarities are typically calculated individually for each category [11].

2. Naïve Bayes classification approach:

It's a simple probability-based prediction referring to Bayes theorem having strong independence assumption (naïve). Previous studies on algorithms of classification have proven that Naïve Bayes is one of the best algorithms in comparison with the other ones such of Decision Tree, Naïve Bayes, and K-NN. It has also been found that accuracy and speed are the most supporting and helpful features of the algorithm in classifying data [12]. The advantage of Bayesian classifier is that it requires small training data set for classification. It is easier for implementation, fast to classify and more efficient. It is non sensitive to irrelevant features. It is used in personal email sorting, document categorization, email spam detection and sentiment detection [13].

3. Decision Tree (DT) classification approach:

Decision trees are trees that classify instances by sorting them based on feature values. Each node in a decision tree represents a feature in an instance to be classified, and each branch represents a value that the node can assume. Instances are classified starting at the root node and sorted based on their feature values. A decision tree, or any learned hypothesis h, is said to over fit training data if another hypothesis h2 exists that has a larger error than h when tested on the training data, but a smaller error than h when tested on the entire dataset. There are two common approaches that decision tree induction algorithms can use to avoid over fitting training data: i) Stop the training algorithm before it reaches a point at which it perfectly fits the training data, ii) Prune the induced decision tree. If the two trees employ the same kind of tests and have the same prediction accuracy, the one with fewer leaves is usually preferred.

Decision trees can be significantly more complex representation for some concepts due to the replication problem. A solution is using an algorithm to implement complex features at nodes in order to avoid replication and speed. One of the most useful characteristics of decision trees is their comprehensibility [14].

IV. EXPERIMENTATION

In this work there are three steps which have done to classify the tweets based on its polarity to positive and negative classes. First, data collection. In this step, we have used the lexicon of Sudanese dialect which have been proposed in [11]. Twitter's API was also used to collect tweets related to Sudanese revolution. The total dataset consists of 6482 tweets with a good balance of positive and negative sentiments.

The second step is data preprocessing. In this step, every tweet was cleaned and removed any aspects from tweets which have no sentiment meaning, such as Numbers (1, 2, ...), URLs and special characters. After that, we have tokenized tweets into words. Then stopwords were removed from tweets and we have removed any word that doesn't have a meaning (articles preposition), such as "في", "على", "من" which mean "in", "on", "from" in English. Also, stemming process have done which includes, removing any affixes (prefixes that added to the beginning of the word, infixes that added to the middle of the word, or/and suffixes that added to the ending of the word) from words to reduce these words to their roots. Then Normalization process have done which includes, eliminating the diacritical markings, non-letters,

letter Hamza (ء). Also, replaced ة and ة with ا, replaced final ي with ي, and replaced final ة with ة.

The output of this steps is a cleaned and ready dataset for labelling to either positive or negative classes.

The third and the last step is classification. In this step the data is divided into training and testing dataset. Training dataset used to build the classification models. And three different classification techniques were used which are SVM, NB and DT to classify the tweets based on its polarity to positive and negative classes.

Table 1,2 below shows an example of tweet in the preprocessing stage and sample of tweets after classification stage.

Table 1: Example of a tweet in preprocessing stage

Preprocessing Step	Tweets After Preprocessing
The original tweet	سقطت ما سقطت صابنهاا ..
Data cleaning	سقطت ما سقطت صابنهاا
Removing Duplicated Characters	سقطت ما سقطت صابنها
Tokenization	سقطت، ما، سقطت، صابنها
Stopword Removal	سقطت سقطت صابنها
Stemming	سقط سقط صب
Normalization	سقط سقط صب

Table 2: Sample of Classified Tweets.

Tweet	In English	Class
البلد خربت	The country is ruined	Negative
الليلة حميدتي قام بالواجب مع اهل السودان	Today Hamidati has done his duty with the people of Sudan	Positive
أجمل ما في الثورة انو شعب السودان اتوحد في مطلبو بسقوط النظام الفاسد	The most beautiful thing in the revolution is that the people of Sudan united in demanding the fall of the corrupt regime	Positive
نزفت دم يا السودان	The blood of Sudan is bleeding	Negative

V. RESULTS

Precision, Recall, Accuracy and F-Measure measures were calculated for every classifier, to evaluate the classification of

tweets as positive or negative. Table 3 below shows a confusion matrix that introduces these measures.

Table 3: Confusion Matrix for Two Classes Pos and Neg

Actual Class	Predicted Class	
	Pos	Neg
Pos	TP	FN
Neg	FP	TN

In the table above we have four parameters (True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN)), which are needed to calculate the measures. **TP**: is the number of tweets that were correctly classified as positive. **TN**: is the number of tweets that were correctly classified as negative. **FP**: is the number of tweets that were classified as positive but they are negative. **FN**: is the number of tweets that were classified as negative but they are positive. Therefore, the formula of the measures is the following:
 Precision = TP/(TP+FP).
 Recall = TP/(TP+FN).
 Accuracy = (TP+TN)/(TP+FP+FN+TN).
 F-measure = 2* (Precision*Recall)/ (Precision+Recall).

In this work cross-validation was performed for evaluate the classification of tweets using SVM classifier with different k-folds (number of folds). Table 4 below, shows the Precision, Recall, Accuracy and F-measure results of the SVM classifiers with different K-folds experiments.

Table 4: The Precision, Recall, Accuracy and F-measure of SVM with K-folds Cross-validation

No of folds	Precision	Recall	Accuracy	F-Measure
K=5	74.3%	96.3%	75.2%	83.9%
K=6	73.2%	95.8%	73.7%	83.0%
K=7	74.2%	96.2%	75.0%	83.8%
K=8	73.8%	95.7%	74.3%	83.3%
K=9	74.0%	95.7%	74.6%	83.5%
K=10	74.2%	95.9%	75.0%	83.8%

From table 4 above we notice that, the best results of Precision, Recall, Accuracy and F-measure were achieved by SVM when used 5-folds Cross-validation (K=5). Figure 1 below shows the comparison between the four measures.

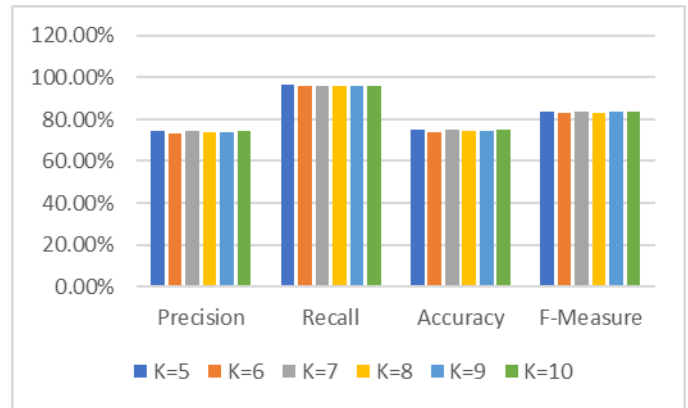


Figure 1: Comparison between the results of SVM based on K-folds cross validation

Table 5 below, shows the Precision, Recall, Accuracy and F-measure results of the NB with different K-folds experiments.

Table 5: The Precision, Recall, Accuracy and F-measure of NB with K-folds Cross-validation

No of folds	Precision	Recall	Accuracy	F-Measure
K=5	86.7%	27.9%	48.7%	42.2%
K=6	85.7%	29.7%	49.4%	44.0%
K=7	88.4%	31.4%	51.1%	46.3%
K=8	87.2%	30.3%	50.2%	44.9%
K=9	86.0%	29.3%	49.4%	43.6%
K=10	86.9%	29.6%	49.7%	44.2%

From table 5 above we notice that, the best results of Precision, Recall, Accuracy and F-measure were achieved by NB when used 7-folds Cross-validation (K=7). Figure 2 below shows the comparison between the four measures.

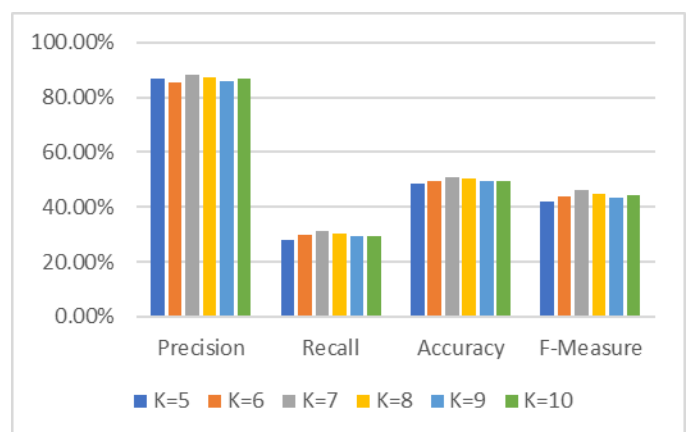


Figure 2: Comparison between the results of NB based on K-folds cross validation

Table 6 below, shows the Precision, Recall, Accuracy and F-measure results of the DT with different K-folds experiments.

Table 6: The Precision, Recall, Accuracy and F-measure of DT with K-folds Cross-validation

No of folds	Precision	Recall	Accuracy	F-Measure
K=5	67.8%	99.8%	68.0%	80.7%
K=6	67.8%	99.6%	68.0%	80.7%
K=7	67.8%	99.8%	68.1%	80.8%
K=8	67.9%	99.9%	68.2%	80.9%
K=9	67.8%	99.8%	68.1%	80.8%
K=10	67.8%	99.8%	68.1%	80.8%

From table 6 above we notice that, the best results of Precision, Recall, Accuracy and F-measure were achieved by DT when used 8-folds Cross-validation (K=8). Figure 3 below shows the comparison between the four measures.

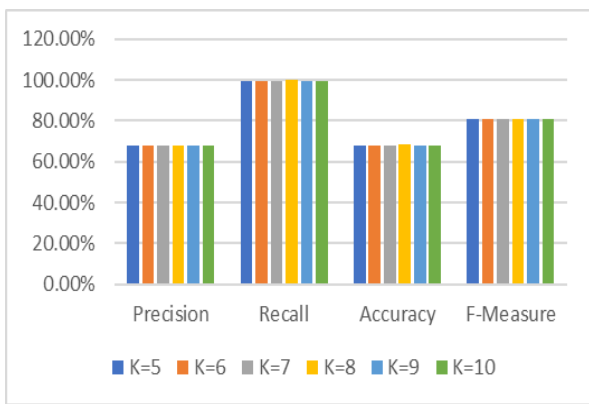


Figure 2: Comparison between the results of DT based on K-folds cross validation

Table 7 below shows a comparison between SVM, NB and DT classifiers based on the best results of Precision, Recall, Accuracy and F-measures.

Table 7: Class Precision, Recall, Accuracy and F-Measure for the SVM, NB and DT classifiers

Classifier	Precision	Recall	Accuracy	F-Measure
SVM	74.3%	96.3%	75.2%	83.9%
NB	88.4%	31.4%	51.1%	46.3%
DT	67.9%	99.9%	68.2%	80.9%

From table 7 above we notice that, Support Vector Machine achieved good results for Accuracy and F-measure which equal to 75.2%, 83.9% respectively. While Naive Bayes achieved good results for precision which equal to 88.41%, and Decision Tree achieved good results for recall which equal to 99.9%.

Figure 4 below shows a composition of the results of the three classifiers in detail.

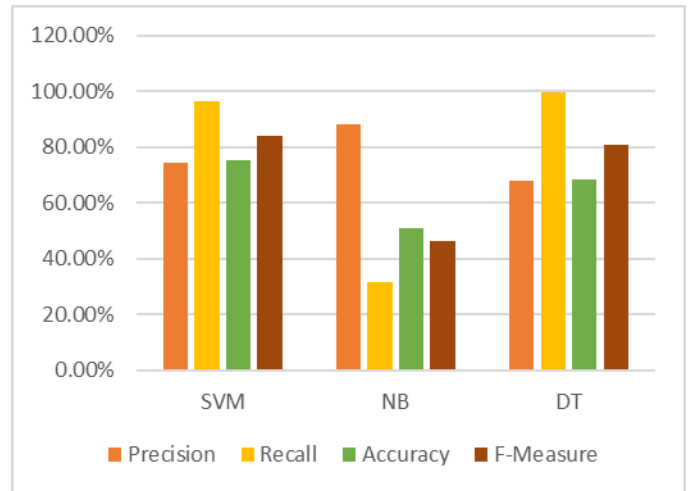


Figure 4: Precision, Recall, Accuracy and F-measure for the SVM, NB, DT classifiers.

From figure 4 above we found that the best Accuracy and F-measure was achieved by SVM. While the best Precision was achieved by NB. In addition, the best Recall was achieved by DT classifier.

VI. CONCLUSION

This paper focused on sentiment classification of twitter data belonging to Sudanese revolution. We have used a lexicon of Sudanese dialect. The total dataset consists of 6482 tweets with a good balance of positive and negative sentiments. It was divided into training and testing sets. The SVM, Naïve Bayes and Decision Tree were applied to detect the polarity of the tweets. The results show that, SVM achieved the best Accuracy and F-Measure and it equals 75.2%, 83.9% respectively. While Naïve Bayes achieved best Precision and it equals 88.41%. Also, the best recall was achieved by Decision Tree and it equals 99.9%.

In addition, based on our data which collected from Twitter, the percentages of positive and negative opinions toward the government was calculated. 9.4% represents the percentage of positive opinions related the government, while 90.6% represents the percentage of negative opinions related the same government.

To the best of our knowledge, the current work is the first to deal with detect the sentiments and classify tweets related to the Sudanese revolution. Also, the percentages of the positive and negative opinions could be very important and valuable for identifying the kind of opinions that the twitter users are sharing. It is also needed to take into account that this is not a sample of the whole Sudanese population but a subset of social networks users.

REFERENCES

- [1] Sadegh, Mohammad, Roliana Ibrahim, and Zulaiha Ali Othman. "Opinion mining and sentiment analysis: A survey." *International Journal of Computers & Technology* 2, no. 3 (2012): 171-178.
- [2] Pradhan, Vidisha M., Jay Vala, and Prem Balani. "A survey on Sentiment Analysis Algorithms for opinion mining." *International Journal of Computer Applications* 133, no. 9 (2016): 7-11.
- [3] Salloum, Said A., et al. "A survey of Arabic text mining." *Intelligent Natural Language Processing: Trends and Applications*. Springer, Cham, 2018. 417-431.
- [4] Koturwar, Praful, Sheetal Girase, and Debajyoti Mukhopadhyay. "A survey of classification techniques in the area of big data." *arXiv preprint arXiv:1503.07477* (2015).
- [5] Jain, Achin, and Vanita Jain. "Sentiment classification of twitter data belonging to renewable energy using machine learning." *Journal of Information and Optimization Sciences* 40.2 (2019): 521-533.
- [6] Qamar, Ali Mustafa, Suliman A. Alsuhibany, and Syed Sohail Ahmed. "Sentiment classification of twitter data belonging to saudi arabian telecommunication companies." *International Journal of Advanced Computer Science and Applications (IJACS)* 1, no. 8 (2017): 395-401.
- [7] Tagwa Abd Elatif Mohammed, "Review of Sentiment Analysis for Classification Arabic Tweets." *International Journal of Emerging Technology and Advanced Engineering*, ISSN 2250-2459 Volume 6, Issue 3, March 2016.
- [8] Al-Subaihini, Afnan A., Hend S. Al-Khalifa, and AbdulMalik S. Al-Salman. "A proposed sentiment analysis tool for modern Arabic using human-based computing." *Proceedings of the 13th International Conference on Information integration and Web-based Applications and Services*. ACM, 2011.
- [9] Al-Kabi, Mohammed N., Amal H. Gigieh, Izzat M. Alsmadi, Heider A. Wahsheh, and Mohamad M. Haidar. "Opinion mining and analysis for Arabic language." *International Journal of Advanced Computer Science and Applications (IJACSA)*, SAI Publisher 5, no. 5: 181-195, 2014.
- [10] Arunadevi, J., S. Ramya, and M. Ramesh Raja. "A study of classification algorithms using RapidMiner." *International Journal of Pure and Applied Mathematics* 119.12 (2018): 15977-15988.
- [11] Huda, J Abdelhameed, and Susana Muñoz - Hernández. "Emotion and opinion retrieval from social media in Arabic language: Survey." *Information and Communication Technologies for Education and Training and International Conference on Computing in Arabic (ICCA-TICET)*, 2017 Joint International Conference on, pp. 1-8. IEEE, 2017.
- [12] Slamet, Cepy, et al. "Web Scraping and Naïve Bayes Classification for Job Search Engine." *IOP Conference Series: Materials Science and Engineering*. Vol. 288. No. 1. IOP Publishing, 2018.
- [13] Rajeswari, R. P., and Kavitha Juliet. "Aradhana. Text Classification for Student Data Set using Naïve Bayes Classifier and KNN Classifier." *International Journal of Computer Trends and Technology (IJCTT)* Vol 43 (2017).
- [14] Elhawary, Mohamed, and Mohamed Elfeky. "Mining Arabic business reviews." *Data Mining Workshops (ICDMW) IEEE International Conference on*. IEEE, 2010.
- [15] Prabhu, BP Aniruddha, B. P. Ashwini, Tarique Anwar Khan, and Arup Das. "Predicting Election Result with Sentimental Analysis Using Twitter Data for Candidate Selection." *In Innovations in Computer Science and Engineering*, pp. 49-55. Springer, Singapore, 2019.
- [16] Zhai, Zhongwu, Hua Xu, Jun Li, and Peifa Jia. "Sentiment classification for Chinese reviews based on key substring features." *In Natural Language Processing and Knowledge Engineering, 2009. NLP-KE 2009. International Conference on*, pp. 1-8. IEEE, 2009.
- [17] Bansal, Barkha, and Sangeet Srivastava. "Lexicon-based Twitter sentiment analysis for vote share prediction using emoji and N-gram features." *International Journal of Web Based Communities* 15, no. 1 (2019): 85-99.
- [18] Ignatow, Gabe, Nicholas Evangelopoulos, and Konstantinos Zougris. "Sentiment Analysis of Polarizing Topics in Social Media: News Site Readers' Comments on the Trayvon Martin Controversy." *Communication and Information Technologies annual (Studies in Media and Communications, Volume 11)* Emerald Group Publishing Limited 11: 259-284, 2016.
- [19] Nakov, Preslav, Sara Rosenthal, Svetlana Kiritchenko, Saif M. Mohammad, Zornitsa Kozareva, Alan Ritter, Veselin Stoyanov, and Xiaodan Zhu. "Developing a successful SemEval task in sentiment analysis of Twitter and other social media texts." *Language Resources and Evaluation* 50, no. 1: 35-65, 2016.
- [20] Soni, Swati, and Aakanksha Sharaff. "Sentiment Analysis of Customer Reviews based on Hidden Markov Model." *Proceedings of the International Conference on Advanced Research in Computer Science Engineering & Technology (ICARCSET 2015)*. ACM, 2015.
- [21] Ceron, Andrea, Luigi Curini, Stefano M. Iacus, and Giuseppe Porro. "Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France." *New Media & Society* 16, no. 2: 340-358, 2014.
- [22] Soliman, Taysir Hassan, M. A. Elmasry, A. Hedar, and M. M. Doss. "Sentiment analysis of Arabic slang comments on Facebook." *International Journal of Computers & Technology* 12, no. 5: 3470-3478, 2014.
- [23] Maynard, Diana, Kalina Bontcheva, and Dominic Rout. "Challenges in developing opinion mining tools for social media." *Proceedings of the NLP can u tag user generated content* 15-22, 2012.
- [24] Asghar, Muhammad Zubair, Fazal Masud Kundi, Shakeel Ahmad, Aurangzeb Khan, and Furqan Khan. "T-SAF: Twitter sentiment analysis framework using a hybrid classification scheme." *Expert Systems* 35, no. 1 (2018): e12233.
- [25] Itani, Maher M., Rached N. Zantout, Lama Hamandi, and Islam Elkabani. "Classifying sentiment in Arabic social networks: Naive search versus naive bayes." *In Advances in Computational Tools for Engineering Applications (ACTEA), 2nd International Conference on*, pp. 192-197. IEEE, 2012.
- [26] Ramanathan, Vallikannu, and T. Meyyappan. "Twitter Text Mining for Sentiment Analysis on People's Feedback about Oman Tourism." *In 2019 4th MEC International Conference on Big Data and Smart City (ICBDSC)*, pp. 1-5. IEEE, 2019.