

# Arabic Tweets Sentiment Analysis about Online Learning during COVID-19 in Saudi Arabia

Asma Althagafi<sup>1</sup>, Ghofran Althobaiti<sup>2</sup>, Hosam Alhakami<sup>3</sup>, Tahani Alsubait<sup>4</sup>  
College of Computer and Information Systems  
Umm Al-Qura University  
Makkah, Saudi Arabia

**Abstract**—The COVID-19 pandemic can be considered as the greatest challenge of our time and is defining and reshaping many aspects of our life such as learning and teaching, especially in the academic year of 2020. While some people could adapt quickly to online learning, others consider it to be inefficient. The re-opening of schools and universities is currently under consideration. However, many experts in many countries suggested that at least one semester should be online, during the pandemic. Understanding the public's emotional reaction to online learning has become significant. This paper studies the attitude of people of Saudi Arabia towards online learning. We have used a collection of Arabic tweets posted in 2020, collected mainly via hashtags that originated in Saudi Arabia. Our sentiment analysis has shown that people have maintained a neutral response to online learning. This study will allow scholars and decision makers to understand the emotional effects of online learning on communities.

**Keywords**—Social media analytics; sentiment analysis; online learning; Arabic tweets

## I. INTRODUCTION

The COVID-19 pandemic has pushed educators towards adopting online learning, starting from the academic year 2020. Looking back at history, the COVID-19 CORONA virus was discovered in the last month of 2019, in Wuhan of China. In March 2020, the Director General of the World Health Organization (WHO)<sup>1</sup> announced COVID-19 as a pandemic, after evaluating the accelerated dissemination and magnitude of the lethal virus across the globe, with an additional declaration of social distancing as a way of halting the spread of the pandemic. This pandemic has prompted a worldwide physical shutdown of businesses, sporting events and schools by forcing all institutions to switch to online channels. As a result, pupils cannot attend schools or institutions physically. Around the same time, though, they need to learn to deal with this condition and to continue their studies. While some people could adapt quickly, others considered it to be inefficient. The re-opening of schools and universities is currently under consideration. However, most experts suggested that at least one semester should be online. So, in this research we explore what people think about online learning via exploring tweets related to online learning to understand people's opinions and attitudes.

People openly share their thoughts and views in no more than 280-characters tweets, making Twitter one of the most popular social networking sites in the world. In this research,

we focus on sentiment analysis of people's posts in Twitter. We argue that tweeting is a good way of raising public opinion about online learning, as the platform is widespread in Saudi Arabia.

At present, sentiment analysis or opinion mining has been considered to be one of the most emerging fields of study sparked by social networks. Sentiment analysis is the job of recognizing optimistic and negative views, feelings, and evaluations. The aim of sentiment analysis is to decide a writer's attitude to some subject or the overall document's tonality [1]. The purpose of Sentiment Analysis is to find views, define the conveyed emotions, then describe their polarity [2]. Sentiment analysis can be conducted at several levels: document level, sentence level and subject level [3]. In this research, we are interested in the sentence level sentiment analysis of Arabic tweets to assess the tweet polarity; whether it is positive, negative or neutral. We are interested in sentiment classification in the Arabic language at the sentence level in which the aim is to classify tweets about online learning in Saudi Arabia to determine people's opinions related to this topic and classify the tweets to positive, negative or neutral.

The remainder of this article is arranged as follows: Section II presents a background to Sentiment Analysis in Arabic. Section III introduces related work. Section IV introduces methods and materials used in this research. Section V exposes our results and discussions. Finally, Section VI lays forth the conclusion and future work.

## II. BACKGROUND

### A. Sentiment Analysis in the Arabic Language

In recent years, social media sentiment analysis has become a hot subject for opinion mining in many social networking applications [4]. Sentiment analysis-based opinion mining may be done by evaluating a subject's feelings and actions about an occurrence or a particular subject. Arabic sentiment analysis is one of the most challenging social media sentiment analysis techniques, owing to the casual noisy content and the rich morphology of the Arabic language. The Arabic opinion analysis approaches are gaining more popularity and significance by rising the rate of feedback and comments by Arabic users on numerous social media platforms [5]. Arabic is a Semitic language that is spoken in the Middle East and North Africa by more than 250 million individuals. It is one of the United States' six official dialects and the language of the Holy Quran. It is additionally the language that a portion of the world's most

<sup>1</sup><https://www.who.int/en>

prominent scholarly, science and chronicled works have been written in. There are three principle types of Arabic [6]:

- 1) Classic Arabic (CA)  
That is the kind of Arabic that the Mushaf (Holy Quran) is written in. The grammar of today's Arabic is significantly different, as Mushaf was written in the 6th century CE. CA is based on the medieval dialects of Arab tribes.  
CA special symbols are used to indicate proper pronunciation and to deliver words. Such written Arabic symbols are almost exclusively found in the Quran or alrecitation [7].
- 2) Modern Standard Arabic (MSA)  
In today's Arabic-speaking countries, it is the most common form of Arabic used. In virtually every media medium, MSA is used in TV, documentaries, papers, and radio broadcasts. Most written papers in seminars and politicians' speeches are in the MSA [6].
- 3) Colloquial Arabic (Arabic)  
It is the Arabic dialect unique to each region, the Arabic language that is utilized to communicate thoughts fundamentally in the WWW, generally in sites, discussions, and conversational posts. While much of its vocabulary and grammatical origins come from the MSA, it still incorporates its own lexicon [6].

### B. Challenges of using Arabic Language in Sentiment Analysis

There are many challenges facing Arabic, some of which are particular to the sentiment analysis activities, and the rest are due to the complexities of the Arabic language. A big challenge is the unavailability of colloquial Arabic sentiment lexicons and the limited availability of MSA lexicons, relative to those constructed in the English language, while most people in social Media platforms use colloquial Arabic to write their opinions and feelings [8]. The use of Latin characters to represent Arabic words is a recent social media theme, which is referred to as Arabizi. Arabic social media users often prefer to switch languages in their writings between Arabic and English, making it impossible to detect whether a phrase written with Latin characters is Arabizi or English [8]. Sarcasm is a kind of speech act in which there's something good a person says when something negative is actually meant, or vice versa [9]. Sarcasm is very difficult to detect, with just a few attempts in English for sarcasm detection using supervised and semi-supervised approaches to learning [9]. No research that deals with sarcasm identification has been found in Arabic sentiment analysis, to the best of our knowledge.

### III. RELATED WORK

Since people and consumers express their thoughts and feelings more freely than ever before, sentiment analysis is becoming an important method for tracking and understanding these feelings. A lot of research has been done on developing methods of sentiment analysis and defining the process of detecting sentiment for different languages around the world.

Heikal et al. [3] examine sentiment analysis of Arabic tweets utilizing Deep Learning. They utilized a troupe model

for investigation. The complexity of the Arabic language has urged them to investigate diverse profound learning models that have not examined to improve the accuracy of Arabic language examination. They use a mix of Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) models. They applied the model to a collection of Arabic tweets. The model accomplishes a 64.46% F1 score, which surpasses the F1 score of the cutting edge profound learning model fusing Convolutionary Neural Network (CNN) and Long Short-Term Memory (LSTM) models, to anticipate a 53.6% inclination on the Arabic Sentiment Tweets Dataset (ASTD).

Aldayel and Azmi [10] present a further investigation on the subject of getting the feeling from Arabic tweets with an attention on those starting in Saudi Arabia. There are numerous hindrances to the act of utilizing dialectical Arabic in tweets; the engendering of spelling mistakes; and the enormous variety of Twitter spaces, to name a few. To depict the extremity of Arabic tweets, they propose a cross breed arrangement that mixes semantic direction with an AI approach. A lexical classifier is utilized to characterize tweets in a solo manner, for example to manage plain tweets, and the SVM classifier further fortifies the result of the lexical classifier. The tests demonstrate that neither the lexical nor SVM classifiers will get the impact of the half breed approach. The normal execution of the crossover classifier regarding F-estimation and accuracy is 84.01% and 84.01% separately, respectively.

Duwairi et al. [11] present another study on Arabic tweets, discussing sentiment analysis. About 350,000 tweets have been collected for this purpose. In order to label 25000+ tweets, crowd sourcing was used. A label like Positive, Negative or Neutral is assigned to each tweet. Majority voting was used with every tweet to determine the final label. There are many novel contributions to this work, such as managing negations, Arabizi and Arabic dialects. The system was tested using three built-in classifiers in Rapid miner. The results obtained are promising.

Zhou et al. [12] present a study about a Tweets Sentiment Analysis Model (TSAM). Their research has shown that it is feasible and can be very useful to develop an intelligent lexicon-based sentiment analysis framework. However the opinion analysis method, in its current form, has not yet achieved its full potential. A variety of study problems must be worked out in order to boost the existing TSAM, such as distinguishing between parts of speech. The opinion terms are extracted in the current model as the features. The accuracy of the part of speech tagging has been found to affect the overall sentiment ratings. Therefore to enhance the existing technique, advanced NLP methods must be implemented. Taking into account the study of emotions, the method identifies positive and negative thoughts, experiences and feelings. It measures a person's feeling in the sense of positivity or negativity. However, the study of text emotions moves beyond the positive-negative dimension to the discreet forms of emotions such as joy, sadness, etc. Moreover, text-based mood analysis, such as text grouping and clustering, poses numerous obstacles beyond conventional text analysis. With the analysis carried out above, the TSAM model will yield even more accurate performance, with the use of more specific entity recognition approaches.

Larkey et al. [13] present one more examination utilizing

an assortment of word references that store positive, negative and neutral roots. A stemmer was used to translate words into roots in order to evaluate the sentiment or class of a sentence. On the chance that the subsequent root shows up in the positive/negative/neutral root word reference, it is called positive/negative/impartial. In the event that the word isn't in the word reference, the user is asked to choose its extremity and afterward to add its source to the relating word reference.

#### A. Comparison of Related Work

Overall, we reviewed several recent studies highly related to ours. Some of these studies, as shown in Table I, applied Sentiment Analysis on tweets to extract the sentiments by using many methods, as we detail in the table.

TABLE I. COMPARISON OF RELATED WORK

Studies	Main Method	Accuracy%
Heikal et al. [3]	Convolutional Neural Network(CNN)	64.46%
Heikal et al. [3]	Long Short-Term Memory(LSTM)	53.6%
Aldayel et al. [10]	Support Vector Machines (SVM)	84.01%

### IV. METHODOLOGY

#### A. Tweets Extraction

Firstly, we got a Twitter Developer Account<sup>2</sup> that helps you to access the Twitter API. By using our API information we collected tweets in Arabic only related to online learning. We searched for the following hashtags in the tweets:

- #التعليم\_عن\_بعد (Online teaching)
- #الدراسة\_عن\_بعد (Online learning)
- #البلاكبورد (Blackboard)
- #تيمز (Teams)
- #التعليم\_الالكتروني (E-Learning)

And the keywords that follow:

- التعليم عن بعد (Online teaching)
- الدراسة عن بعد (Online learning)
- البلاكبورد (Blackboard)
- تيمز (Teams)
- التعليم الالكتروني (E-Learning)

By using the get\_tweets function, we got 10445 tweets that also include venue, username, retweet count, favorite count, and time of tweet.

#### B. Data Pre-Processing

The first stage of pre-processing is getting rid of duplicate tweets. There were 2269 duplicate tweets, so the number of tweets was decreased to 8176. Then, we need to get rid of stop words, punctuation, hashtags, comparisons, links, and one or two-letter words as follow:

- Tokenize tweets into words and punctuation marks. The sentence "البلاكبورد اليوم معلق." can be tokenized like [".", "معلق.", "اليوم", "البلاكبورد"]
- Remove URLs from tweets because the URLs are pointed to extra information that was not a prerequisite for sentiment analysis in our approach. We also removed numbers, punctuation marks and extra white spaces because they do not contain emotions.
- Remove stop words. Removing stop words from text helps to recognize the most relevant words. Here we delete terms like: (who, whom, whose, not) الذي, التي, التي, الذي by using the stop words from the nltk library.
- Finally we apply stemming which is a natural language processing technique that solves the issue of vocabulary mismatch [14] and keeps only the origin of each word.

#### C. Sentiment Analysis

Once the tweets were pre-processed, the second stage is sentiment analysis. In this step, we can focus on our main aim in this project which is to measure sentimental characteristics of tweets, such as polarity and subjectivity, using TextBlob<sup>3</sup>. Polarity is a variation in value between '-1' and '1'. It shows us how positive or negative the statement is. Subjectivity is another difference of value between '0' and '1' which shows whether the statement is an opinion or statement. Textblob comes with the core features of natural-language processing essentials; this approach classifies the polarity of textual data in positive, neutral and negative groups as '1', '0' and '-1'. We divided sentiments into three groups, namely positive, negative and neutral, based on their polarity, as seen in Table II.

TABLE II. POLARITY CLASSIFICATION

Value of Polarity	Sentiment
>0	Positive
0	Neutral
<0	Negative

#### D. Evaluation

To evaluate our results, we split our data into 80% train and 20% test sets. To classify the texts, a variety of machine learning algorithms have been implemented. We used the Naive Bayes, Random Forest and K-nearest neighbor Classifiers.

<sup>2</sup><https://developer.twitter.com/en>

<sup>3</sup><https://textblob.readthedocs.io/en/dev/>

Naive Bayes is a simple but fast classification algorithm. It is a commonly used algorithm for classifying documents [15]. Multinomial classifier Naive Bayes is widely used in the case of text categorization. It depends on three assumptions: documents are generated by a mixture model, between each mixture component and class there is a one-to-one correspondence, and each mixture component is a multinomial distribution of terms [16]. Theorem of Bayes offers a way to measure the posterior likelihood by equation 1.

$$P(c | x) = \frac{P(x | c)P(c)}{P(x)} \quad (1)$$

Where:

- $P(c | x)$  is the posterior probability of the attribute given by the class.
- $P(c)$  is a prior probability of class.
- $P(x)$  is a prior probability of attribute.
- $P(x | c)$  is the probability of attribute given class.

Random Forest (RF) classifiers excel in a number of automated sorting functions, such as categorization and emotion analysis. It is ideal for treating high dimensional noise data in text classification [17]. The phases of the Random Forest algorithm are as follows: The first step is to conclude that  $n$  samples and  $T$  classification attributes are found in the training set.  $N$  samples are collected using the bootstrap sampling process to get a new sample collection. In the second step, the  $t(T \leq t)$  attributes are selected at random from the  $t$  attributes given. The optimal classification node is chosen by using the optimal feature norm of a decision tree such that all the sub samples are leaf nodes. Repeat the second step of  $K$  in the third step, create  $K$  decision trees, and get the final random forest. In the fourth step, the function model of the classifier is  $H(x)$ , the decision tree is  $h_i$ , the classification label is  $y$ , and the indicator function is  $I(h_i(x) = Y)$ . The random-forest decision-making formula is as equation 2 [18]:

$$H(x) = \operatorname{argmax} \sum_{i=1}^k I(h_i(x) = Y) \quad (2)$$

K-nearest neighbor (KNN) is a standard example based classifier that does not make a clear, declarative description of the category, but depends on the category labels attached to the training documents identical to the test text [19]. KNN is classified by a majority vote of its neighbors, with the case assigned to the most common class of its nearest K-neighbors, determined by a distance function in equation 3. For  $K = 1$ , the case is simply assigned to the class of its closest neighbor.

$$\sqrt{\sum_{i=1}^k (X_i - Y_i)^2} \quad (3)$$

## V. RESULTS AND DISCUSSIONS

To have a better understanding of the public opinion towards online learning, we studied the sentiment people expressed in social media in the first academic term of 2020 in Saudi Arabia by tweets in the Twitter platform. We got the

results as Fig. 1 shows. That most tweets were expressing a neutral sentiment that might have happened because most of the tweets contained sentences that do not express negative, or positive emotions, the rest are due to the complexities of the Arabic language, such as having no sentiment lexicon available for colloquial Arabic, while MSA lexicons are limited relative to those constructed for the English language, and most peoples in social Media platforms use colloquial Arabic to write their opinions and feelings.

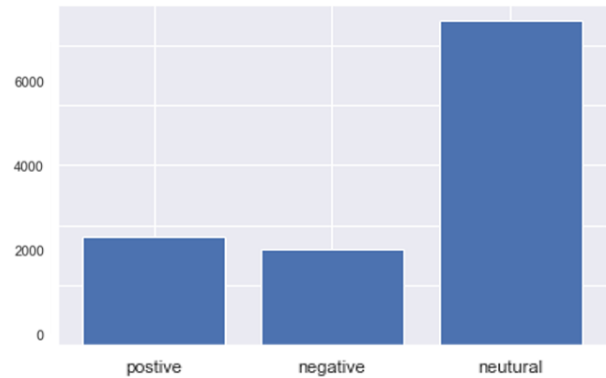


Fig. 1. Distribution of Sentiment.

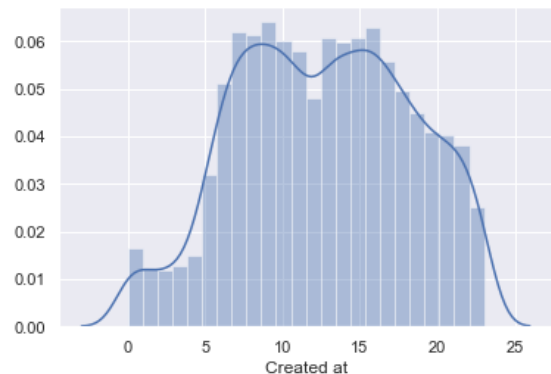


Fig. 2. Hourly Distribution of Tweets.

We also obtained the hourly distribution of tweets as shown in Fig. 2. We can see that the amount of tweets during the day is increased in the period from 6am until 3pm. This period has the most activity because these times are the times of lectures and courses for students in Saudi Arabia.

Moreover, after using TextBlob we can get words clouds for each label. Let us look at the positive and negative tweets words as seen in Fig. 3 and 4. Apparently, in negative words, people whose tweets are negative finds online learning tedious, terrible, and stressful. On the other hand, some positive people prefer online learning opportunities.

To evaluate our results the algorithm performance review experimental environment is supported by the Windows 10 operating system, Intel(R) Core(TM) i7-4710HQ CPU 2.50

## VI. CONCLUSION

Our research was centered on Twitter's opinion mining and sentiment analysis about online learning during COVID-19 pandemic, which bifurcates tweets based on three categories: positive, negative and neutral. Our goal was to get a better understanding of the feelings and opinions of tweeters about online learning. To do so, we collected about 10445 tweets. After that we applied sentiment analysis to these tweets and measured sentimental characteristics of tweets, such as polarity and subjectivity, using TextBlob. We got that most tweets were expressing a neutral sentiment, that might have happened because most of the tweets contained sentences that does not express negative nor positive emotions. One of the key challenges, however, is the lack of resources to be able to analyze the Arabic language, especially that each country has different colloquial Arabic. As for future work, we plan to understand people's attitudes towards different platforms of online learning via sentiment analysis of the feelings shared by the public about these platforms.



Fig. 3. Positive WordCloud.



Fig. 4. Negative WordCloud.

GHz cpu and 16.0 GB memory as support for the whole experiment and we used Python as the language of programming. We used a machine learning model by Naive Bayes, RF and KNN Classifiers. We split our data into 80% train and 20% test sets. We got 77% of Naive Bayes, 84% of RF and 67% of KNN Classifier.

The comparison of Naive Bayes, RF and KNN classifiers for multi-class text classification is also presented in this research. The findings indicate that the RF multi-class classification method achieved the highest classification accuracy in comparison with Naive Bayes and KNN classifiers because it works well with high-dimensional data such as a text classification compared to the other classifiers model.

## REFERENCES

- [1] A. Shouky and A. Rafea, "Sentence-level arabic sentiment analysis," in *2012 International Conference on Collaboration Technologies and Systems (CTS)*, pp. 546–550, 2012.
- [2] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams engineering journal*, vol. 5, no. 4, pp. 1093–1113, 2014.
- [3] M. Heikal, M. Torki, and N. El-Makky, "Sentiment analysis of arabic tweets using deep learning," *Procedia Computer Science*, vol. 142, pp. 114–122, 2018.
- [4] S. Ahmad, M. Z. Asghar, F. M. Alotaibi, and I. Awan, "Detection and classification of social media-based extremist affiliations using sentiment analysis techniques," *Human-centric Computing and Information Sciences*, vol. 9, no. 1, p. 24, 2019.
- [5] H. AlSalman, "An improved approach for sentiment analysis of arabic tweets in twitter social media," in *2020 3rd International Conference on Computer Applications & Information Security (ICCAIS)*, pp. 1–4, IEEE, 2020.
- [6] E. H. Mohamed and E. M. Shokry, "Qsst: A quranic semantic search tool based on word embedding," *Journal of King Saud University-Computer and Information Sciences*, 2020.
- [7] A. A. Khrisat and Z. A. Alharthy, "Arabic dialects and classical arabic language," *Advances in Social Sciences Research Journal*, vol. 2, no. 3, 2015.
- [8] N. Al-Twairsh, H. Al-Khalifa, and A. Al-Salman, "Subjectivity and sentiment analysis of arabic: trends and challenges," in *2014 IEEE/ACS 11th International Conference on Computer Systems and Applications (AICCSA)*, pp. 148–155, IEEE, 2014.
- [9] B. Liu, "Sentiment analysis and opinion mining," *Synthesis lectures on human language technologies*, vol. 5, no. 1, pp. 1–167, 2012.
- [10] H. K. Aldayel and A. M. Azmi, "Arabic tweets sentiment analysis—a hybrid scheme," *Journal of Information Science*, vol. 42, no. 6, pp. 782–797, 2016.
- [11] R. M. Duwairi, R. Marji, N. Sha'ban, and S. Rushaidat, "Sentiment analysis in arabic tweets," in *2014 5th International Conference on Information and Communication Systems (ICICS)*, pp. 1–6, IEEE, 2014.
- [12] X. Zhou, X. Tao, J. Yong, and Z. Yang, "Sentiment analysis on tweets for social events," in *Proceedings of the 2013 IEEE 17th international conference on computer supported cooperative work in design (CSCWD)*, pp. 557–562, IEEE, 2013.
- [13] N. Farra, E. Challita, R. Abou Assi, and H. Hajj, "Sentence-level and document-level sentiment mining for arabic texts," in *2010 IEEE international conference on data mining workshops*, pp. 1114–1119, IEEE, 2010.

- [14] L. S. Larkey, L. Ballesteros, and M. E. Connell, "Improving stemming for arabic information retrieval: light stemming and co-occurrence analysis," in *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 275–282, 2002.
- [15] P. Melville, W. Gryc, and R. D. Lawrence, "Sentiment analysis of blogs by combining lexical knowledge with text classification," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1275–1284, 2009.
- [16] K. P. Nigam, "Using unlabeled data to improve text classification," tech. rep., CARNEGIE-MELLON UNIV PITTSBURGH PA SCHOOL OF COMPUTER SCIENCE, 2001.
- [17] T. Salles, L. Rocha, and M. Gonçalves, "A bias-variance analysis of state-of-the-art random forest text classifiers," *Advances in Data Analysis and Classification*, pp. 1–27, 2020.
- [18] Y. Sun, Y. Li, Q. Zeng, and Y. Bian, "Application research of text classification based on random forest algorithm," in *2020 3rd International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE)*, pp. 370–374, IEEE, 2020.
- [19] S. Tan and J. Zhang, "An empirical study of sentiment analysis for chinese documents," *Expert Systems with applications*, vol. 34, no. 4, pp. 2622–2629, 2008.