

ASPECTS DETECTION MODEL FOR USERS' REVIEWS USING MACHINE LEARNING TECHNIQUES

SAAD MOAWAD KHALIFA¹, MOHAMED IBRAHIM MARIE², MAI MAHMOUD EL-DEFRAWI³

^{1,2,3} Information Systems dept, Faculty of Computers and Artificial Intelligence, Helwan University, Cairo, Egypt

E-mail: ¹ khalifasaad43@fci.helwan.edu.eg, ² mohamedmarie@yahoo.com,

³ mai.eldefrawi@fci.helwan.edu.eg

ABSTRACT

Over time, sentiment analysis and opinion mining emerged as significant study areas. In order to determine a person's mood from a written text, sentiment analysis examines texts, feelings, and views and divides them into positive and negative categories. Sentiment analysis is used on social media sites, where enormous amounts of data are created daily. Fewer studies have been conducted on the Arabic language compared to those conducted on the English language because of how difficult and complicated sentiment analysis is in the morphologically complex Arabic language. The Coronavirus pandemic was one of the factors that increased the amount of research being published in the area of sentiment analysis, which was then utilized to identify and categorize people's emotions throughout the Coronavirus era. In order to categorize a group of brief texts about employees' opinions on working from home during the coronavirus pandemic into binary positive and negative feelings and to identify the challenging aspects, or the issues that employees face while working from home, so that business owners can review and address them, machine learning algorithms were used in this study. Additionally, other lengthy writings on the same subject were subjected to the same algorithms so that we could monitor the model's progress while we worked on both short and long texts. With the aim of determining if text length may affect efficiency and accuracy, it was discovered that working with lengthy texts increased accuracy. The machine learning techniques Logistic Regression (LR), Random Forest (RF), Multinomial Naive Bayes (MNB), and Support Vector Classification (SVC) have been chosen and used. According to the results, MNB performs best when dealing with short texts, outperforming the competition with an accuracy rate of 88.5 %. It is followed by LR, RF, and SVC, which each achieve accuracy rates of 87.6 %, 78.1 %, and 87.6 %. SVC wins the competition when applied to lengthy texts, with an accuracy rate of 98.5 %, followed by LR, MNB, and RF, with accuracy rates of 97.6 %, 96.6 %, and 97.5 %, respectively.

Keywords: *Sentiment Analysis, Machine Learning, Opinion Mining, Social Media, Coronavirus.*

1. INTRODUCTION

Recently, the artificial intelligence sector has taken an interest in the domains of sentiment analysis and opinion mining, which are concerned with evaluating texts, feelings, and moods of individuals and categorizing them as positive or negative. One factor that contributed to this is the daily generation of a large amount of data through social media platforms like Facebook and Twitter, which individuals use to communicate their feelings, emotions, and attitudes by posting about their present moods and the challenges they face [1].

Twitter has over 300 million active users each month who use it to express situations and opinions. Some people use social media to share their

experiences with a service or product, either positively or negatively. Through this information, businesses can be aware of this information, which is considered feedback, and take it into account to increase features and improve defects. Many people spend the majority of their time on social media platforms and find it to be a safe place to publish and share their diaries in Modern Standard Arabic (MSA) which is the standard shape of Arabic language or in his dialect which is an informal shape of the language. Some governments also use this information. People use social media networks to express their interests, feelings, and learn new things. Satisfaction surveys are conducted [2-4].

To determine whether an opinion is good or negative, Natural Language Processing (NLP)

techniques which describe the interaction between computer and human language like sentiment analysis and opinion mining are utilized. Due to the Arabic language's morphological complexity and a large number of dialects it comprises, the majority of published study has been done on the English language [5].

In addition to being the language of the Holy Quran, Arabic is one of the Semitic languages, spoken by over 300 million people worldwide. MSA and Arabic dialects are now the two most widely used dialects of the Arabic language. Arabic speakers make up 180 million internet users, or 4.8 % of all users worldwide [3],[6].

Sentiment analysis based on the Arabic language is considered one of the most challenging techniques due to the informal and noisy content as well as the Arabic language is morphologically complex, this field attracted researchers the continuous rising of averages of people's comments and feedback by Arabic users comfortable share their opinions and thoughts social media platforms, for example, Twitter which is considered is one of the most popular and widespread platforms which many people used to share diaries in less than 280 characters. The objective of sentiment analysis and opinion mining is to detect the writer's point of view about some subject. Sentiment analysis can be done on some levels: document level, sentence level, and word level. Covid-19 is considered one of the more challenging issues in real-time many of our concepts of life have changed such as learning and working whereas with the exacerbation of the pandemic working and learning become online and from home and that made it even more important to measure the public emotions, attitude, and opinions about the changes that the Coronavirus has forced upon us [7].

It is normal to tweet about daily events, situations, and social issues. For instance, consider the significant role that Twitter played during the Egyptian Revolution, when Arabic users used Twitter on average at a rate of 40 to 45 tweets per minute. As a result, twitter is thought to be fertile ground to gather opinions. There are currently about 11 million active users on Twitter, and the number of tweets has surpassed the 849 million marks [8].

Covid-19 was discovered in Wuhan City in December 2019. By 28 February and 31 March, there were 858,361 persons infected with the illness, and 210 nations had been affected. Many of these countries had made the decision to shut down their nation. People were compelled to stay at home due to the lockdown. Work-from-home policies have been implemented by several firms [9].

Due to the various social media platforms, many employees were sent home by their companies in order to safeguard them from the deadly Coronavirus outbreak. Employees used to voice their ideas regarding the modern work and living styles [10].

WFH is an acronym for “work from home” which means that work is done from home instead of the office. The acronym “WFH” is used as a nickname for the concept.

Covid-19 pushed many people to work from home and the opinions differ between supporters and opponents.

In addition to identifying the aspects or issues that the employees are upset about so that businesses and decision-makers can improve the working environment for those who work from home, this research introduces a classification model for categorizing and predicting the opinions of Arabic people about this issue while doing so during the COVID-19 pandemic using TF-IDF which is considered one of the most popular and commonly used term-weighting schemes. Selected machine learning algorithms to accomplish this task were SVC which is a type of deep learning algorithms that makes a supervised learning for regression and classification, MNB which is a popular learning approach in NLP based on Bayesian theorem, RF which is used for classification and regression using ensemble method, LR which is a supervised learning statistical method to predict binary outcome.

It also examines the impact of the length of the text or opinion on the classification process. Texts whose length is less than or equal to the maximum length of the tweet which is 288 characters long were treated as short texts, and the texts whose length exceeded the maximum length of the tweet were treated as long texts.

The remaining sections of the essay are structured as follows: The related work pertinent to the study's goal—machine learning-based sentiment analysis of Arabic texts—and the difficulties with Arabic-language social media material are shown in Section 2. The proposed model is presented in Section 3. The experiment, findings, qualitative comparison, and features are all described in Section 4. The major conclusions of this study and next research are described in Section 5.

2. RELATED WORK

This section briefly describes the related works in the domain of sentiment analysis related to analyzing long and short Arabic text.

The purpose of Victor Ivan López Rodriguez study [10], was to respond to the serious

issue, "How do individuals feel about WFH?" Twitter data was used as a dataset during the Coronavirus Pandemic, and the method used to obtain this Twitter data involved first creating a Twitter developer account, from which one obtained the access credentials, followed by the use of the Tweepy library to extract data; retweets were excluded from this research work. NLP techniques were used to work with the data and clean it, including tokenization, stop words removal, stemming, and lemmatization. After that, it was ready to be used to detect sentiment using several libraries, including "TextBlob" and "seaborn, and matplotlib" for plotting the processed data. The findings show that the majority of individuals are content working remotely during the COVID-19 epidemic.

Budoor Al-Marzooqi, et al. [11], the objective of this study was to survey employees' opinions about WFH experience the number of the hypothesis was 3, the first was about the strongly satisfied worker about this experience, the second was about workers find WFH is an advantageous thing, third was about opinions that believe that will differ by gender. The methodology used was the Quantitative approach, through making a questionnaire containing twenty-four multiple choice questions, to ensure the correctness of the content of the report, the opinion of two experts was requested setting Cronbach Alpha at 0.737 with a total of 219 responses. The results showed that most of the opinions are satisfied with this experience.

Akash D Dubey et al. [12], conducted a study to measure the sentiments of Work-from-Home among the employee using Twitter data using Twitter API for the collection of the tweets, the search terms were #WorkFromHome or #WFH like what was done in [10], RTweet package was used in R to avoid the duplication of the tweets, retweets were excluded, After data collection process they went through a data preprocessing phases which contain tokenization, stop words removal after that they used NRC Emotion Lexicon (Mohammad SM, Turney PD, 2013) was used for the sentiment analysis to analyze the Tweets From the analysis of the sentiments regarding Work-from-Home. It was found that 73.10 % of the tweets had positive sentiments as compared to 26.10% negative sentiments.

Abdulrahman Radaideh et al. [13], the objective was to focus on the usage of specific apps in the UAE such as Zoom, Totok, and Botim for conducting meetings over the internet. Twitter data was used as a dataset and collected almost 833 tweets using Twitter API and converted the image to text after data collection they went through preprocessing phase

after that they built two models to detect the real-time sentiment analysis. Naïve Bayes (NB) and Recurrent Neural Network (RNN) were used. The results showed that NB is more accurate with (84%) than (RNN) with (79%).

Asma Althagafi et al. [14], the objective of this was about e-learning in Saudi Arabia during Covid-19 pandemic, dataset used was extracted from Twitter preprocessing steps were applied were tokenization, removing URLs, numbers, white spaces, punctuations, stop word and stemming, for machine learning algorithms that have been applied were NB, RF, K-nearest neighbor (KNN). the results showed that Random Forest outperforms other algorithms.

Abeer A. AlFutamani et al. [15], conducted a study that focuses on the extraction of emotions that is related to tweets that were written in Arabic Saudi Dialect, with the help of the Supervised Learning Approach that is concerned with the categorization of the words mentioned in the text into emotional categories in addition to using two datasets in their analysis first was from Twitter and second was with the help and facilities of STC. In the step of preprocessing URLs, email addresses, links, ads, pictures, and other media, irrelevant information, hashtags, retweets, punctuation, special characters, and repeated letters were removed in addition to normalization. In the step of feature extraction, N-gram was applied, Algorithms that were applied MNB, Support Vector Machine (SVM), LR, and the results of algorithms were remarkably close.

Islam Saif Eldin Mukhtar Heamida et al. [16], conducted a study that focus on analyzing Facebook comments that were written in colloquial and modern Sudanese language whose number has reached 1050 comments addition they created a lexicon that contains one thousand words because there is no a ready-made dictionary for Sudanese colloquial dialect and words were binary classified as a positive or negative word. In the data preprocessing Phase duplicate words or characters, stop words, URLs, tags, duplicate characters that show emotion, and characters that were written in the English language were removed. In addition to tokenization, normalization and stemming techniques were applied. Algorithms that have been applied were SVM and Naïve Bayes (NB) the results showed that the SVM Algorithm outperform NB Algorithm.

Dhiaa A. Musleh et al. [17], conducted a study that focus on developing a model for analyzing tweets of Arab users in addition to detecting depression. The dataset was collected from Arab

users that use Twitter who answered a survey called CES-D which is a brief survey that contains twenty questions that were designed to measure depression symptoms. In the preprocessing, URLs, hashtags, punctuation, usernames, repeated letters, Arabic diacritics, and stop words were removed as well as applying Tokenization and Stemming in the phase of feature extraction TF-IDF was applied. Applied Algorithms were RF and KNN. RF obtained the highest accuracy and the KNN showed the least accuracy.

Belgacem Brahim et al. [18], conducted a study that focus on analyzing and classifying movie reviews that were written in the Arabic language using a supervised learning approach at the document level they considered the whole document as a unit of information. OCA dataset was employed (Opinion Corpus for Arabic) which is commonly used in research studies and the number of reviews has reached five hundred. In the preprocessing step, they removed unrelated parts of reviews, tokenizing, removing punctuation, keeping Arabic words only as well as removing non-Arabic letters, normalization, and removing diacritics. They applied N-gram for the feature extraction step. They applied the SVM algorithm for classification.

Abinash Tripathy et al. [19], conducted a study that focus on sentiment classification using a hybrid machine learning approach at the document level using two datasets that contain movie reviews called IMDb and polarity in which movie reviews were written in the English language. In the preprocessing step stop, words and irrelevant information were removed, and they applied a combination of SVM for selecting the best features in the training datasets to be an input to the RNN. Some studies stated a theoretical comparison between short and long text in the classification process

M. Ali Fauzi et al. [20], sentiment analysis in short Casual texts such as product reviews is a more complex and challenging task. Short texts are noisy, sparse, and suffer from a shortage of context information, unlike long text. The solution that is presented is necessary to convert short text to long text by enriching it with semantics to be like a large document.

Analyzing short texts is a complex and challenging task because of the shortness of the text which may be a reason for ambiguity reasons due to the lack of more information that it has [21].

Short texts such as tweets are a complex and challenging task because texts might contain a

combination of formal and casual language, emojis, symbols, and special characters. Therefore, it is an often-difficult task to understand the context of the text and it is complex to extract the suitable emotions expressed by users [2].

This paper focus on classifying employees' opinions about WFH during coronavirus pandemic and identifying challenging problems that affect their productivity as well as measuring the impact of text length on the models is the main concern of the research.

2.1 Challenges Associated with Arabic Content on Social Media

This subsection will briefly illustrate challenges associated with Arabic content spread on social media.

There are a few challenges associated with Arabic content on social media to summarize:

- Unstructured and casual text.
- Spelling errors.
- Colloquial words.
- Frequent use of abbreviations.
- Tending to repeat letters to represent feelings.

Arabic verbs may take multiple forms like "حب" may appear in different many forms like "أحب", "يحب", "تحب", "تحب", "يحبوا", "يحبوا". In addition, some Arabic names, especially family names, are derived from adjectives, which can be confusing and deceptive in the sentiment analysis process, which can be overcome by using Part-of-speech (POS) tagging through which this word can be known as a noun or an adjective but this solution is difficult to apply in the case of working on dialects like "سعيد" which may be adjective or noun and that may be confusing. The use of diacritics in the Arabic language is also confusing thing because words of the same root can give different and contradictory emotional impressions like "يلعب", "تلاعب" these two words have the same root "لعب" but the sentiment of "يلعب" is positive and the sentiment of "تلاعب" is negative, another example is "تميز", "امتياز" these two words have the same root "ميز" but the sentiment of "امتياز" is positive and the sentiment of "تميز" is negative [23].

3. PROPOSED MODEL

The proposed model of this research paper is applying sentiment analysis on two datasets that

take about opinions of employees about working during the Covid-19 pandemic which is the first dataset contains short texts opinions and the second dataset contains long texts opinions. The sentiment analysis proposed model starts with the data collection process, data preprocessing, feature extraction, classification, and evaluation. The proposed model is shown in figure 1.

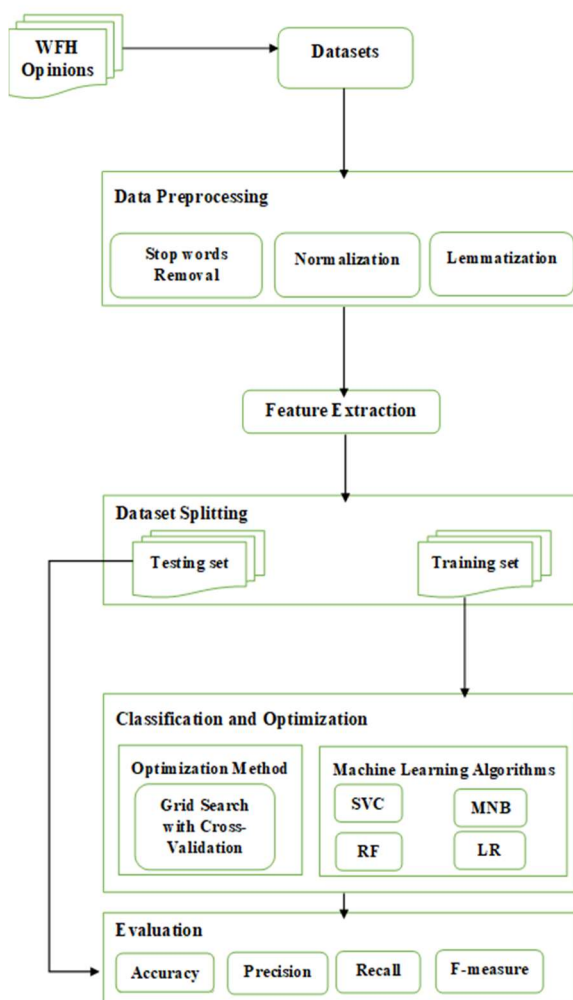


Figure 1: Sentiment Analysis Proposed Model

A. Data Collection

In this research, our concern was focusing on texts and opinions which were interested in working on the topic during the covid-19 pandemic two datasets used in this research that was written in Modern Standard Arabic (MSA) and other Arabic dialects. The first one contains short text opinions and the second one contains opinions with long text opinions and two datasets were talking about the same topic, each dataset contains 1046 opinion but the length of opinions in each dataset are different. If

we look at a sample of the research that was posted in the sentiment analysis field such as Budoor Al-Marzooqi, et al. [11], who worked on responses collected from a distributed questionnaire with twenty-four multiple choice questions and the number of responses reached 219.

Abdulrahman Radaideh et al. [13], worked on 833 tweets that discussed the usage of specific apps in UAE such as Zoom, Took, and Botim for conducting meetings over the internet. Islam Saif Eldin Mukhtar Heamida et al. [16], worked on 1050 comments that focused on analyzing Facebook comments written in colloquial and modern Sudanese language.

Lamiaa Mostafa [24], worked on sentiments collected from approximately one thousand students in the English language using a google sheet. After we examined a sample of some research in the field of sentiment analysis, we will find that the number of samples that will be worked on in this research is considered on average. For the first dataset, which contains short textual opinions, we will review the methods and sources that were used to collect it, Twitter (60 %) Using Twitter API, an Online survey (10 %), Translated tweets from English (20%) so that they are in line with our Arabic language and the behavior of our Arab society in addition to Opinions (Short Text) of Arab and English People that have been spread Online (10 %).

For the second dataset that contains opinions with long texts, methods and sources that were used to collect it were YouTube Comments and videos Description That talking about Issues (20 %), Facebook Posts and comments (20 %), news and topics from the online digital press (40 %), Translated news and topics from English Press and Opinions (Long Text) of Arab and English People that have been spread Online (20 %). All these opinions in two datasets were manually annotated by three experts in the Arabic language into (Positive or Negative) opinions. Table 1 contains the characteristics of the datasets, and figure 2 contains a sample of opinions in each dataset.

Table 1: Characteristics of Datasets.

Opinions	Datasets	
	Long	Short
Total	1046	1046
Negative (Neg)	523	523
Positive (Pos)	523	523

D. Classification and Machine Learning Algorithms

There are fewer different machine learning algorithms currently being used for sentiment classification and prediction tasks. Logistic Regression, a standard algorithm, is used for binary classification or problems with two label values many fields use Logistic Regression including medical, and marketing fields, and it focuses on prediction and explanation. Support Vector Machine (SVM) is a linear algorithm, and its idea is to determine the best border that separates two classes of data, in addition to its usage for classification and regression according to a number of researches SVM outperforms other algorithms when used in sentiment analysis. RF is considered one of the most popular algorithms and widely used in data science it consists of a vast number of individual trees that works as one group each tree is considered to predict a class and the class with a higher number of predictions becomes the winner model among other prediction models [25].

Naïve Bayesian is a probabilistic method whose base is Bayes' theorem. There are many forums from this algorithm such as MNB which has been used in this study that is used as a baseline for the classification of text it works with a document as a group of words and considers information and word frequency [26].

In this research paper, the utilized algorithms are Support Vector Classification (SVC), Random Forest Classifier (RF), Multinomial Naïve Bayes (MNB), and Logistic Regression (LR). Each dataset is divided into a training set to train the model and a testing set to evaluate it in the same proportion.

E. Evaluation

All these algorithms were evaluated for each dataset by a group of classification metrics such as precision, recall, and F-measure. The functions of Precision, Recall, Accuracy, and F-measure (F1) evaluation metrics are shown in (1), (2), (3), and (4) respectively.

$$\text{Precision (P)} = \frac{(TP)}{(TP+FP)} \quad (1)$$

$$\text{Recall (R)} = \frac{(TP)}{(TP+FN)} \quad (2)$$

$$\text{Accuracy (Acc)} = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (3)$$

$$\text{F-measure} = \frac{(2*precision*recall)}{(precision+recall)} \quad (4)$$

Where (TP) true positive, (TN) true negative, (FP) false positive, and (FN) false negative.

A **true positive** is an outcome where the model correctly predicts the positive class. Similarly, a **true negative** is an outcome where the model correctly predicts the negative class. A **false positive** is an outcome where the model incorrectly predicts the positive class.

4. EXPERIMENT AND RESULTS

This section will present the results associated with applied algorithms MNB, LR, RF, and SVC along with the two datasets when the test size (T) equals 10% and 20%.

4.1 Short Text Classification

Machine learning algorithms have been applied to this dataset with short opinions such as SVC, LR, RF, and MNB and these algorithms have been evaluated by a group of evaluation metrics such as accuracy, precision, recall, f-measure, and it was found that Multinomial Naïve Bayes (MNB) outperforms other algorithms, Table 2 shows an explanation for all evaluation metrics that associated with short text dataset.

Table 2: Short Text Classification

Algorithm	Label	T (%)	P (%)	R (%)	Acc (%)	F1 (%)
LR	Neg	10	94	83	87.6	88
			81	94		
	Pos	20	92	80	84.7	85
			77	91		
MNB	Neg	10	93	86	88.5	89
			84	91		
	Pos	20	89	84	85.2	85
			80	87		
SVC	Neg	10	92	84	87.6	88
			83	91		
	Pos	20	90	79	83.3	83
			76	89		
RF	Neg	10	83	76	78.1	78
			73	81		
	Pos	20	81	72	74.2	74
			68	77		

4.2 Long Text Classification

The same machine learning algorithms, the same preprocessing techniques, and the same feature extraction techniques in addition to the same

splitting size of the short text dataset have been applied to the dataset with long text and it is also evaluated in terms of accuracy, precision, recall, and f-measure, In the following Table 3 an explanation for all evaluation metrics associated with long text.

Table 3: Long Text Classification.

Algorithm	Label	T (%)	P (%)	R (%)	Acc (%)	F1 (%)
LR	Neg	10	97	98	97.1	97
	Pos		98	96		
	Neg	20	97	98	97.6	98
	Pos		98	97		
MNB	Neg	10	96	96	95.2	95
	Pos		94	96		
	Neg	20	97	96	96.6	97
	Pos		96	97		
SVC	Neg	10	97	98	97.1	97
	Pos		98	96		
	Neg	20	97	100	98.5	99
	Pos		100	97		
RF	Neg	10	98	98	98.1	98
	Pos		98	98		
	Neg	20	97	98	97.6	98
	Pos		98	97		

4.3 Qualitative Comparison

The objective of this research in general is to enrich the research in sentiment analysis using Arabic language because the challenging nature of Arabic language makes research in this area much more difficult compared to what was published using English language, and the contribution of this research especially focus on Classifying employees' opinions about WFH during covid-19 and identifying problems they face, Measuring the impact of text length (short & long) on the model accuracy.

In this subsection, we briefly compare our proposed model and some of the related works, Table 4 shows the comparison in the case of short text and Table 5 for long text.

Table 4: Qualitative Comparison of the Short Text.

Models	Dataset	Langu-age	Technique
Proposed	A group of opinions was compiled from various sources with a total of 1046 samples.	Arabic and its dialects .	SVC, MNB, LR, RF
Related [10]	A distributed questionnaire with a total of 219 responses.	Arabic.	Quantitative approach.
Related [11]	Twitter dataset with a total of 2000 tweets.	English .	TextBlob, Naïve Bayes.
Related [14]	Twitter dataset with a total of 8176 tweets.	Arabic, Saudi Arabia dialects .	KNN, NB, RF.

Table 5: Qualitative Comparison of the Long Text.

Models	Dataset	Langu-age	Techniques
Proposed	A group of opinions was compiled from various sources with a total of 1046 samples.	Arabic and its dialects	SVC, MNB, LR, RF.
Related [18]	Movie reviews (OCA) with a total of 500 reviews dataset.	Arabic.	SVM
Related [19]	Two different movie reviews datasets are IMDb (50,000) reviews and polarity (2000) reviews.	English .	SVM RNN

From tables 4 and 5, we can say that it is not fair for the comparison to be fair between our model and these previous studies due to the difference in sample size, the various sources from which samples were collected, the difference in the language of the samples that were worked on. In addition to the different techniques used in text processing. For example, in some research it stemming techniques were applied but, in our research, we decided to use lemmatization instead. In addition to the differences in used approaches whether were machine learning, deep learning, quantitative approach, etc.

4.4 Aspects

This subsection introduces one of our main goals which is to detect problems that employees face during working from home during the covid-19 pandemic.

The Wordcloud method is a generator that is used to capture words that frequently occur in your dataset, through which you can get more valuable information that helps you to know people's feelings about something. Wordcloud has been used in [10], to capture words that frequently occur whether positive or negative and mentioned that is a guide for many companies and employers to know the feeling of their employees about working from home during the pandemic.

In this study, wordcloud has been used to capture aspects that are associated with negative sentiment because it is considered a problem because it has been repeated by many people in our dataset and to be able to capture the most relevant aspects of Part of Speech (POS) tags to select only words, adjectives and nouns that stand out the meaning, not including such stop words that don't have an effect on the sentiment analysis or text mining. Figure 3 shows wordcloud that contains aspects concerning negative sentiment.



Figure 3: Aspects Associated with Negative Sentiment.

In figure 3. We will find words such as "المفاصل", "الرقية", and "الاجور" these words for example are considered a key for decision making and employers to improve productivity during working from home during a covid-19 pandemic by Providing employees with comfortable chairs and creating a healthy work environment in addition to wage increase here we tried to give a magical help for employers through standing out some problems that face the employee to control it and increasing their productivity.

Authors Contributions: Formal analysis, Mohamed Marie; Investigation, Mai Mahmoud; Methodology, Saad Moawad; Software, Saad Moawad and Mohamed Marie; Supervision, Mohamed Marie; Validation, Mai Mahmoud;

Writing-original draft, Mai Mahmoud. All authors have read and agreed to the published version of the manuscript.

5. CONCLUSION

In this research paper, our goal was to build a classification model, to treat a critical issue that most people face now, which is what are the opinions of the employees about WFH during the covid-19 pandemic, and what are the problems they face to help the employee work well and in a comfortable environment, in addition, it helps decision-makers and business owners to increase their productivity by solving these problems that their employees face.

The scarcity of data and no ready-made dataset that contain a huge volume of data associated with the topic of the research was a challenging task.

With the help of machine learning algorithms named SVC, RF, LR, and MNB we became able to achieve our all planned goals, algorithms have been applied to a dataset containing a group of short opinions it was found that MNB algorithms outperform other algorithms in the case of short text and its f-measure was 89 % with a test size of 10 %, to measure the impact of the text length on the model accuracy the same algorithms have been applied to another dataset with long text, RF showed efficiency when dealing with the long text, it achieved accuracy reaching 98 % with a test size of 10 %, SVC showed efficiency with a test size of 20 %.

We conclude from this that the text length has effects on the efficiency of the model, As the length of the text increases, the accuracy of the model increases, because the model has a large amount of information and data to learn from it. We also conclude from this that, the longer text is, the greater the similarity between opinions and each other, which leads to focusing on the words that stand out in the sentiment, as there is similarity and the repetition of some important words that have an impact on the classification process.

The nature of the text and the number of records in the dataset can be the main reasons to make the results of the models vary, this work focused on Arabic text only in addition to the size of the dataset wasn't enormous.

As a feature work, we are planning to make a comparison between the accuracies obtained from machine learning algorithms to the accuracies that will be obtained from deep learning algorithms with further data preprocessing techniques to improve the classification process.

REFERENCES:

- [1] A. Alsayat and N. Elmitwally, "A Comprehensive Study for Arabic Sentiment Analysis Challenges and Applications", *Egyptian Informatics Journal*, Vol. 21, No. 1, 2020, pp.7-12.
- [2] A. Aloqaily, M. Al-Hassan, K. Salah, B. Elshqeir, and M. Almashagbah, "Sentiment Analysis for Arabic Tweets Datasets Lexicon-Based and Machine Learning Approaches", *J. Theor. Appl. Inf. Technol.*, Vol. 98, No. 4, 2020, pp. 612-623.
- [3] O. Oueslati, E. Cambria, M.B. HajHmida, and H. Ounelli, "A Review of Sentiment Analysis Research in Arabic Language", *Future Generation Computer Systems*, Vol. 112, 2020, pp.408-430.
- [4] A.A. Al Shamsi and S. Abdallah, "Text Mining Techniques for Sentiment Analysis of Arabic Dialects Literature Review", *Adv. Sci. Technol. Eng. Syst. J.*, Vol. 6, No. 1, 2021, pp.1012-1023.
- [5] O. Al-Harbi, "A Comparative Study of Feature Selection Methods for Dialectal Arabic Sentiment Classification using Support Vector Machine", Vol. 19, No. 2, 2019, pp.167-176.
- [6] I. Awajan and M. Mohamad, "A Review on Sentiment Analysis in Arabic using Document Level", *Int. J. Eng. Technol.*, Vol. 7, No. 3.13, 2018, pp.128-132.
- [7] A. Althagafi, G. Althobaiti, H. Alhakami, and T. Alsubait, "Arabic Tweets Sentiment Analysis about Online Learning during COVID-19 in Saudi Arabia", *International Journal of Advanced Computer Science and Applications*, Vol. 12, No. 3, 2021, pp.620-625.
- [8] M. Zayed, H. Mousa, and M. Elmenshawy, "Sentiment Analysis for Arabic Social Media IJCI", *International Journal of Computers and Information*, Vol. 7, No. 1, 2020, pp.14-31.
- [9] A.D. Dubey and S. Tripathi, "Analysing the Sentiments towards Work from Home Experience during Covid-19 Pandemic", *Journal of Innovation Management*, Vol. 8, No. 1, 2020, pp.13-19.
- [10] Victor Iván, "Home Office due to COVID-19 - Twitter Data Analysis", 2020.
- [11] B. Almarzooqi and F.A. Alaamer, "Worker Perceptions on Working from Home during COVID-19 Pandemic Empirical Evidence from Bahrain", *Available at SSRN 3643890*, 2020.
- [12] A.D. Dubey, "Twitter Sentiment Analysis during COVID-19 Outbreak", *Available at SSRN 3572023*, 2020.
- [13] A. Radaideh, F. Dweiri, and M. Obaidat, "A Novel Approach to Predict the Real Time Sentimental Analysis by Naive Bayes & RNN Algorithm during the Covid Pandemic in UAE", In *2020 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI)*, 2020, pp. 1-5.
- [14] A. Althagafi, G. Althobaiti, H. Alhakami, and T. Alsubait, "Arabic Tweets Sentiment Analysis about Online Learning during COVID-19 in Saudi Arabia", *International Journal of Advanced Computer Science and Applications*, Vol. 12, No. 3, 2021, pp. 620-625.
- [15] A.A. AlFutamani and H.H. Al-Baity, "Emotional Analysis of Arabic Saudi Dialect Tweets Using a Supervised Learning Approach" *Intelligence Automation and Soft Computing*, Vol. 29, No. 1, 2021, pp.89-109.
- [16] I. Saif Eldin Mukhtar Heamida and A.L. Samani Abd Elmutalib Ahmed, "The Classification Model Sentiment Analysis of the Sudanese Dialect Used Into the Internet Service in Sudan", In *Enabling Machine Learning Applications in Data Science*, 2021, pp. 369-378.
- [17] D.A. Musleh, T.A. Alkhales, R.A. Almakki, S.E. Alnajim, S.K. Almarshad, Alhasania, *et al.*, "Twitter Arabic Sentiment Analysis to Detect Depression Using Machine Learning", *CMC-Computers Materials & Continua*, Vol. 71, No. 2, 2022, pp.3463-3477.
- [18] B. Brahim, M. Touahria, and A. Tari, "Improving Arabic Sentiment Classification using a Combined Approach", *Computación y Sistemas*, Vol. 24, No. 4, 2020, pp.1403-1414.
- [19] A. Tripathy, A. Anand, and S.K. Rath, "Document-Level Sentiment Classification using Hybrid Machine Learning Approach", *Knowledge and Information Systems*, Vol. 53, No. 3, 2017, pp.805-831.
- [20] M.A. Fauzi, N. Firmansyah, and T. Afrianto, "Improving Sentiment Analysis of Short Informal Indonesian Product Reviews using Synonym Based Feature Expansion", vol. 16, no. 3, 2018, pp. 1345-1350.
- [21] S. Alosaimi, M. Alharthi, K. Alghamdi, T. Alsubait, and T. Alqurashi, "Sentiment Analysis of Arabic Reviews for Saudi Hotels using Unsupervised Machine Learning", *Journal of Computer Science*, Vol. 16, No. 9, 2020, pp.1258-1267.

-
- [22] G. Alwakid, T. Osman, and T. Hughes-Roberts, "Challenges in Sentiment Analysis for Arabic Social Networks", *Procedia Computer Science*, Vol. 117, 2017, pp.89-100.
- [23] L. Mostafa, "Egyptian Student Sentiment Analysis using Word2vec during the Coronavirus (Covid-19) Pandemic", In *International Conference on Advanced Intelligent Systems and Informatics*, 2020, pp. 195-203.
- [24] A.A. Sayed, E. Elgeldawi, A.M. Zaki, and A.R. Galal, "Sentiment Analysis for Arabic Reviews using Machine Learning Classification Algorithms", In *2020 International Conference on Innovative Trends in Communication and Computer Engineering (ITCE)*, 2020, pp. 56-63.
- [25] M. Abbas, K.A. Memon, A.A. Jamali, S. Memon, and A. Ahmed, "Multinomial Naive Bayes Classification Model for Sentiment Analysis", *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, Vol. 19, No. 3, 2019, pp.62-67.