



## Emotion Detection in Arabic Text using Machine Learning Methods

By Fatimah Khalil Aljwari

*University of Jeddah*

**Abstract-** Emotions are essential to any or all languages and are notoriously challenging to grasp. While numerous studies discussing the recognition of emotion in English, Arabic emotion recognition research remains in its early stages. The textual data with embedded emotions has increased considerably with the Internet and social networking platforms. This study aims to tackle the challenging problem of emotion detection in Arabic text. Recent studies found that dialect diversity and morpho- logical complexity in the Arabic language, with the limited access of annotated training datasets for Arabic emotions, pose the foremost significant challenges to Arabic emotion detection. Social media is becoming a more popular kind of communication where users can share their thoughts and express emotions like joy, sadness, anger, surprise, hate, fear, so on some range of subjects in ways they'd not typically neutralize person. Social media also present different challenges which include spelling mistakes, new slang, and incorrect use of grammar. The previous few years have seen a giant increase in interest in text emotion detection. The study of Arabic emotions might be a results of the Arab world's considerable influence on global politics and thus the economy. There are numerous uses for the automated recognition of emotions within the textual content on Facebook and Twitter, including company development, program design, content generation, and emergency response. in line with recent studies, it's possible to identify emotions in English-language information.

**Keywords:** emotion detection, machine learning, arabic text, KNN, DT, SVM, naive bayes.

**GJCST-G Classification:** FOR Code: 170203



*Strictly as per the compliance and regulations of:*



# Emotion Detection in Arabic Text using Machine Learning Methods

Fatimah Khalil Aljwari

**Abstract-** Emotions are essential to any or all languages and are notoriously challenging to grasp. While numerous studies discussing the recognition of emotion in English, Arabic emotion recognition research remains in its early stages. The textual data with embedded emotions has increased considerably with the Internet and social networking platforms. This study aims to tackle the challenging problem of emotion detection in Arabic text. Recent studies found that dialect diversity and morpho- logical complexity in the Arabic language, with the limited access of annotated training datasets for Arabic emotions, pose the foremost significant challenges to Arabic emotion detection. Social media is becoming a more popular kind of communication where users can share their thoughts and express emotions like joy, sadness, anger, surprise, hate, fear, so on some range of subjects in ways they'd not typically neutralize person. Social media also present different challenges which include spelling mistakes, new slang, and incorrect use of grammar. The previous few years have seen a giant increase in interest in text emotion detection. The study of Arabic emotions might be a results of the Arab world's considerable influence on global politics and thus the economy. There are numerous uses for the automated recognition of emotions within the textual content on Facebook and Twitter, including company development, program design, content generation, and emergency response. in line with recent studies, it's possible to identify emotions in English-language information. "However, we are tuned in to only some initiatives to include Arabic content. Hence, we shall develop a machine- learning model for emotion detection from Arabic textual data on social platforms. This study categorizes the texts supported emotions, anger, joy, sadness, and fear, using supervised machine learning approaches. We used five different machine learning algorithms, namely Decision Tree (DT), K-Nearest Neighbor (KNN), Naive Bayes (NB), Multinomial Naive Bayes (NB), and Support Vector Machine (SVM) to classify emotions in Arabic tweets. These algorithms assessed our proposed approach on the dataset of Arabic tweets provided by SemEval-2018 for EI-oc. This meant that the results of the machine learning approaches were admissible. These results found that the selection Tree and K- Nearest Neighbor classifiers have the simplest accomplishment regarding accuracy, 0.74, While the NB and Multinomial NB classifiers acquired 0.69, and also the SVM obtained 0.63.

**Keywords:** emotion detection, machine learning, arabic text, KNN, DT, SVM, naive bayes.

**Author:** Computer Science and Artificial Intelligence Department  
University of Jeddah, Jeddah, Saudia Arabia.  
e-mail: fatima6794o5@gmail.com

## I. INTRODUCTION

Currently, social media plays a necessary role in way of life and practice. immeasurable individuals use social media for various purposes. Every second, a major amount of knowledge flow via online networks, containing valuable information that may be extracted if the information are correctly processed and analyzed[8]. Social networking media became essential for expressing emotions to the planet due to the fast growth of the web. Several individuals use textual content, audio, video, and images to express their emotions or perceptions[9]. The Affective Computing research field has been an energetic research domain and recently gained great popularity. It aims at providing machines with a human-like ability to grasp and answer human emotions, with more natural interaction between humans and machines[10]. Emotions are a vital component of human life. Emotions affect human decision- making and can enable us to speak with the planet in a very better way. Emotion detection, also called emotion recognition, identifies an individual's feelings or emotions, for instance, joy, sadness, or fury [9]. "Emotion detection," "emotion analysis," and "emotion identification" are all expressions that are periodically used interchangeably. The sentiment analysis could be a means of evaluating if data is negative, positive, or neutral. In contrast, emotion recognition specifies different human emotion types, like joy, love, sadness, happiness, anger, and surprise[9]. quite 400 million people speak Arabic, the official language of twenty-two countries. It is the Internet's fourth most generally used language [10]. Languages utilized in social media, such as Twitter, differ wildly from that utilized on other platforms, like Wikipedia. The English language has been highly determined within the emotion detection field, including datasets and dictionary availability, in contrast to the Arabic, which has minimal resources[11]. Emotion analysis has different applications in every aspect of our existence, including making efficient e-learning frameworks in step with the emotion of scholars, enhancing human-computer interactions, observing the mental state of people, enhancing business strategies supported customer emotions, analyzing public feelings on any national, international, or the political event, identifying potential criminals by analyzing the emotions of individuals after an attack or crime, and improving the

performance of chatbots and other automatic feedback frameworks[4]. In Additionally, text emotion analysis has been a promising research topic. Analyzing the texts and identifying emotion from the words and semantics could be a difficult challenge. The paper aims to automatic recognition of emotions in texts written in the Arabic language by employing a model for Emotion Classification (EC) into emotion classes: Sadness, joy, fear, and anger with the algorithms of machine learning. This approach utilized the SemEval- 2018 Task1 reference dataset and focused on four emotion classes (Joy, Sad, Angry, and fear). Five forms of algorithms are used supported the machine learning approach, namely K-Nearest Neighbor (KNN), Decision Tree (DT), Support Vector Machine (SVM), Naive Bayes (NB), and Multinomial (NB). KNN, DT, SVM, NB, and Multinomial NB classifiers are utilized in the classification process since they offer the foremost satisfactory and better accuracy results among all other classifiers. The findings showed that the choice Tree and K-Nearest Neighbor classifiers have the best accomplishment regarding accuracy, 0.74, While the NB and Multinomial NB classifiers acquired 0.69, and also the SVM obtained 0.63. The structure of this study continues to section II, which presents Problem Definition and Algorithm while section III offers the recent related work on Arabic emotion recognition. Section IV describes the Methodology and results of emotion analysis from Arabic texts and discusses the results. Section V provides the conclusion and future work.

## II. PROBLEM DEFINITION AND ALGORITHM

This section briefly presents the Problem Definition and Algorithm for the detection of emotion in texts written in the Arabic language.

### a) Problem Definition

Most research papers in this field focus on negative or positive emotion analysis and do not go deeply into emotion analysis, especially in Arabic. Research in emotion analysis for Arabic has been minimal compared to other languages like English. This paper addresses the emotion detection problem in Arabic tweets and presents a model to categorize emotions into sadness, joy, anger, and fear. Furthermore, the current work can provide many benefits for governments, health authorities, and decision-makers to monitor people's emotions on social media content. Additionally, it can improve business strategies according to customers' emotions and recognize potential criminals when analyzing people's emotions after an attack or crime.

### b) Algorithm Definition

The Machine Learning approach learns from the info and tries to hunt out the relation between a given

input text and also the corresponding output emotion by building a prediction model. This approach is split into two categories:

#### i. Supervised learning approach

Based on a labeled or annotated dataset, the supervised approach takes a component of this data for the training process using an emotion classifier. This trained data is then examined, and a model is made. The remaining data within the dataset is classed supported this previously trained classifier into the emotion category.

#### ii. Unsupervised learning approach

The unsupervised approach relies on a non-labeled dataset. The approach inherent the drawbacks of the ML algorithm. It requires an oversized dataset for the training process to be accurate. The Machine Learning approach solves the emotion detection issue by categorizing texts into various emotion classifications using the mentioned algorithms. This process is usually done employing a supervised or unsupervised ML technique. To categorize the tweet into each categorization (anger, joy, sadness, and fear), we applied five different supervised machine-learning approaches: KNN, DT, SVM, NB, and NB. Following could even be an inventory of the classifiers discussed during this work:

##### a. K-Nearest Neighbor (KNN)

KNN is addition- ally a fairly AI supported machine learning algorithms in classification, processing, statistical pattern recognition, and much of more. This method in our experiment can classify an emotion correctly[15]. KNN classifies a replacement instance within the test set supported the shortest distance between it and numeric neighbors (k) stored within the training set using the Euclidian Distance equation[13].

##### b. Decision Tree (DT)

A Decision Tree could even be a mode of a tree structure utilized in classification and regression models. It breaks down the datasets into smaller subsets and incrementally develops them into nodes and leaves. The branches of the selection tree represent the category of the datasets. the selection tree is split into four emotion classes: joy, sadness, anger, and fear. The selection tree's goal is to substantiate it achieves maximum separation among classes at each level.

##### c. Support Vector Machine (SVM)

SVM could even be a supervised ML algorithm. The model is straightforward, and far of individuals value more highly to use this model thanks to its less computational power and it gives significant accuracy. SVM conducts linear classification and performs non-linear classification alright [6]. This model's idea is straightforward: The algorithm plots each data item as some extent in n-dimensional space representing the

number of features. Then, it'll use hyper-plane to differentiate between features and classes of emotion.

d. *NB*

NB includes several algorithms of classification based on the Bayes Theorem. The NB classifier presents significant results when it is used for text analyzing data. Such an algorithm offers a prospect examining the study's dataset [12].

e. *Multinomial (NB)*

Multinomial NB classifier works on the concept of term frequency, which suggests what percentage times the word occurs during an extremely document. MNB is specially designed for text data and a particular version of Naive Bayes [6]. MNB tells two facts about whether the word appears during a very document and its frequency there in document.

### III. LITERATURE REVIEW

Although there are many studies during this domain, one amongst the tough challenges for all researchers during this domain is to use emotion analysis and classification for Arabic tweets, which remains limited, most Arabic studies specialise in sentiment analysis to classify tweets into positive/negative classes, underestimating the utilization of emotion detection and analysis to draw down different emotions. The literature review presents the foremost recent works on emotion detection in languages.

Mansy, A et al.[1] researchers proposed an ensemble deep learning approach to research Emotion from user text in Arabic Tweets. They evaluated using the SemEval-2018-Task1- dataset published in a very multilabel classification task. The proposed model was supported three deep learning models. Two models are particular styles of Recurrent Neural Networks (RNNs), the Bidirectional Gated Recurrent Unit (Bi-GRU) and Bidirectional Long Short Term Memory Model (Bi-LSTM). The third may be a pretrained model (PLM) supported Bidirectional Encoder Representations from Transformers (BERT) NAMED MARBERT. The results of the proposed ensemble model showed outperformance over the individual models (Bi- LSTM, Bi-GRU, and MARBERT). They showed an accuracy of 0.54, precision of 0.63, 0.55 in an exceedingly recall, 0.70 in Macro F1 Score, and 0.52 in micro F1 Score.

In addition, Khalil et al. [2] proposed a Bi-LSTM deep learning model for EC in tweets written in Arabic that were employed in the SemEval-2018 dataset. The Aravec with CBOW for the word embedding phase has been employed in feature extraction. Their results have shown an Accuracy of 0.498, and a Micro F1 score of 0.615.

Another study on Arabic emotion analysis was proposed in [3]. The authors addressed the emotion detection problem in Arabic tweets. A tweet may have

multiple emotional states (for example, joy, love, and optimism). during this case, the emotional classification of tweets is framed as a multilabel classification problem. The proposed approach combined the transformer-based Arabic (AraBERT) model and an attention- based LSTM-BiLSTM deep model. The approach used a publicly available benchmark dataset of SemEval-2018 Task 1, where the dataset is formed for multilabel detection of emotion in these tweets. The findings show that such an approach presents accuracy of nearly 54.

A multilabel classification was employed to detect emotions in Arabic tweets by [4]. The authors proposed three models: the Deep Feature-based (DF), the Human engineered feature- based (HEF) model, and the Hybrid model of both models (HEF and DF). They assessed the execution of the proposed model on the SemEval-2018, IAEDS, and AETD datasets. For feature extraction, they used Hourglass of emotions, frequency-inverse document, Lexical sentiment features, and Lexical emotion features. the most effective performance results for the hybrid model were achieved when combining the TF-IDF of unigrams, TF-IDF of the Part of Speech (POS) tags, HGE, LSF, and LEF with the DF model. The findings report that the hybrid model exceeded the HEF and DF models in the datasets. The hybrid achieved for every IAEDS, AETD, and SemEval-2018 datasets an accuracy of 87.2, 0.718, and 0.512, respectively.



Ref no.	Authors\Year	language	Approach	Detection method	Emotion model	Features	Dataset	Evaluation metric and Results
[1]	Mansy, A et al.,2022	Arabic	Deep Learning approach	Bi-LSTM, Bi-GRU, and MARBERT	Arabic Multilabel Emotions Classification	Word Embedding features	SemEval -2018- Task1- Ar-Ec dataset	Accuracy: 0.540 Macro F1 Score: 0.701 Precision: 0.634 Recall: 0.550 Micro F1 Score: 0.527
[2]	Khalil, E. A. H et al.,2021	Arabic	Deep Learning approach	BiLSTM	Multilabel Arabic Emotions Classification model	Word Embedding features	SemEval -2018- Task1 dataset	Validation accuracy=0.575%
[3]	Elfaik et al.,2021	Arabic	Deep Learning approach	LSTM-BiLSTM	Anger,anticipation, disgust, fear, joy, love, optimism, pessimism, sadness, surprise, and trust	Embedding features	SemEval -2018 dataset	Accuracy =(53.82%)
[4]	Alsawaidan el al.,2020	Arabic	Hybrid approach	Deep feature based(DF)model and Human engineered feature based(HEF)model	Plutchik's model+love,optimism, and pessimism emotions	Stylistic,lexical, syntactic, and semantic features.	SemEval -2018 dataset	Accuracy, HEF 44.8% DF 50.5% Hybrid(HEF+DF) 51.20%
[5]	AlZoubi et al., 2020	Arabic	Ensemble and Deep learning approach	Bidirectional GRU_CNN, conventional neural networks,and XGBoost regressor	Anger, fear joy,and sadness	TF-IDF, word-level embedding, and lexicon features	SemEval -2018 dataset	Precision , Ensemble=(69.%)
[6]	Hussein, A et al., 2020	Arabic	Machine learning approach	Naïve Bays, k-Nearest Neighbors (KNN), and Support Vector Machine (SVM).	Happy, Sad, Angry, fear	Word Embedding, TF-IDF features	Text Mining Data	Accuracy, NB =70%, SVM =68.33%, KNN=51.67%
[7]	Saad, M. M et al.,2018	Malay language	Machine learning approach	Support Vector Machine (SVM) and Decision Tree (DT)	Four types of emotions Happy, angry, fearful) and sad)	Term Frequency-Inverse Document Frequency (TF-IDF)	200 Malay children short stories datasets	Accuracy, SVM=30.0%, DT=62.5%

Fig. 1: Summary of the emotion detection in Text

Furthermore, deep learning detects emotions. For instance, AlZoubi et al. [5] have implemented an ensemble approach that contains Conventional Neural Networks (CNN), Bidirectional GRU-CNN (BiGRU-CNN), and XGBoost regressor (XGB) to be utilized in solving the EC of the SemEval-2018 dataset written within the Semitic. The ensemble approach used TF-IDF, word-level embedding, and lexicon features. Results show that their model achieved a precision of 69. In addition, Hussein, et al. [6] followed the machine learning method to detect emotion in Text Mining Data supported Arabic Text. They collected text mining data from the internet while focusing on four emotion classes (sad, happy, afraid, and angry). Three sorts of techniques are used supported machine learning approaches include KNN, NB, and SVM algorithm. The findings also showed that

NB had the best accomplishment regarding accuracy. NB classifiers achieved 70, comparing to SVM that obtained 68.33, while KNN yielded 51.67. Saad et al. [7] proposed a similar model to categorize emotions from the Malay language. The dataset used consists of Malay children's short stories. over 200 short stories were collected, each story varying from 20-50 words. The TF-IDF is extracted from the text and classified using SVM and DT. Four common emotions, happy, angry, fearful, and sad, are classified using the 2 classifiers. Results showed that the choice Tree outperformed the SVM by a 22.2 accuracy rate.

Fig 1 summarizes the acceptable emotion methods reviewed during this paper and sorted on the most recent date.

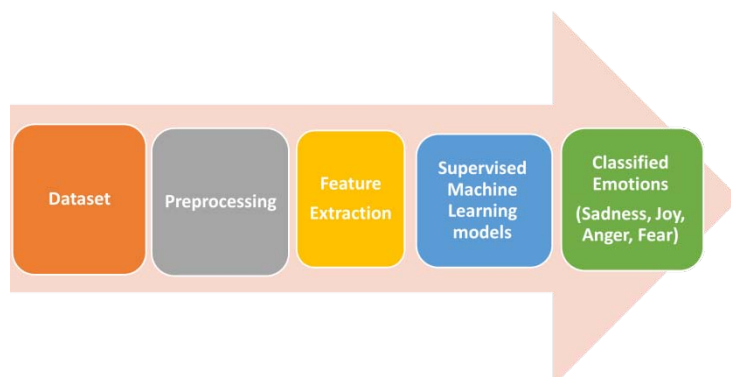


Fig. 2: Shows the main phases of the methodology

ID	Tweet	Affect	Dimension	Intensity Class
1	2018-Ar-00759 @mnwry1 أنتي ساقطه ! ش	anger	3: high amount of anger can be inferred	
3	2018-Ar-03121 ... فعل #قطر من غدر وخيانة ذكرنا بفعل صدام الأسود	anger	3: high amount of anger can be inferred	
10	2018-Ar-04092 ..فصيل قاسم وقناه الجزيره سبب خراب سوريا ربنا ين	anger	3: high amount of anger can be inferred	
17	2018-Ar-03289 @Al3mriRami @HolyIviuss كل مافي الأمر اني غاض	anger	3: high amount of anger can be inferred	
28	2018-Ar-02472 @mm_888s ... انا ضد هذا الكلام كل واحد ظروفه ليش	anger	3: high amount of anger can be inferred	
...	...	...	...	
854	2018-Ar-03946 ..القلب ما يسكنه غير الهم و بوح و بكى #الهم من ا	sadness	3: high amount of sadness can be inferred	
856	2018-Ar-03206 ..أنا إنسان #ضائع هنا أتحدث عن أوجاعي وربما سأتح	sadness	3: high amount of sadness can be inferred	
865	2018-Ar-00640 أوحش أنواع #الزعل انك تكون زعلان علي #نفسك	sadness	3: high amount of sadness can be inferred	
886	2018-Ar-02781 ... يقتلني الحنين ليك حتى تعجز عيوني ذرف آخر دموع	sadness	3: high amount of sadness can be inferred	
888	2018-Ar-01328 ..الوجع الحقيقي هو الذي يفضحك من صوتك من غصة الو	sadness	3: high amount of sadness can be inferred	

Fig. 3: Snapshot of the training dataset before preprocessing

ID	Tweet	Affect	Dimension	Intensity	Class
5	2018-Ar-01214 ...روحي غير حافظتك ي مجرمه اتقوه عليكى عميله وسخ	anger	3: high amount of anger can be inferred		
7	2018-Ar-01154 فليتقوا غضب الحليم إذا غضب	anger	3: high amount of anger can be inferred		
15	2018-Ar-04217 ...والله انها بصمت عار علي جبين الامه العربيه الذ	anger	3: high amount of anger can be inferred		
16	2018-Ar-03520 ... امي: ميخالف*اخي: *يصرخ على امي يسبها يزفها	anger	3: high amount of anger can be inferred		
19	2018-Ar-03153 ... يشوش على ايش الخيانه عبادة ايران خيانة الجار	anger	3: high amount of anger can be inferred		
...	...	...	...	...	...
333	2018-Ar-04365 ... شعور حزين : اذا كنت تضحك وفجأة تتذكر و تسكت -	sadness	3: high amount of sadness can be inferred		
353	2018-Ar-01900 ...إن قلبي ليتمزق تمزقا متى تذكرت أن عيني لا تملك	sadness	3: high amount of sadness can be inferred		
361	2018-Ar-01312 ...ومن "خيبة الـ \n والكلام \n، من كثر المعائب اتعبت	sadness	3: high amount of sadness can be inferred		
365	2018-Ar-04247 ...انا حزين و ما أخفي عليك حزني وأنت داري ان غياب	sadness	3: high amount of sadness can be inferred		
366	2018-Ar-01398 ...ياقلبي المسكين لاحد يحس بغمرة وجعك الدفين جمر	sadness	3: high amount of sadness can be inferred		

Fig. 4: Snapshot of the testing dataset before preprocessing

#### IV. EXPERIMENTAL EVALUATION

This section discusses the used dataset and the methodologies to identify emotions of tweets written in Arabic by utilizing (5) algorithms of machine learning: KNN, DT, SVM, NB, and Multinomial NB. The final section highlights the outcomes of this process.

The main phases of the methodology are shown in Fig. 2. The methodology consisted of the dataset, preprocessing, features engineering, supervised machine learning, and classified emotions based on four emotions (anger, joy, sadness, and fear).

##### a) Dataset

This section discusses the used dataset in which the experiments are performed using the reference emotion detection SemEval-2018 (Affect in Tweets) dataset. The dataset is the public benchmark dataset created for the detection of emotions in tweets written in Arabic. Each tweet is labeled as one of the emotions ( joy, anger, sadness, and fear. ). All these

tweets are in Arabic text. We used only the Eloc for our experiment with four basic emotion categories. The training dataset trains the classifier and the test dataset examines the structured model to show identify the value of trained model. Figure (3) highlights the training dataset and Figure (4) highlights testing dataset.

##### b) Preprocessing

Data preprocessing is taken into account one among the essential phases in machine learning to avoid misleading results and obtain better insights. during this section, the preprocessing steps are discussed as follow: The tweet from the SemEval2018 dataset has been preprocessed using the foremost common preprocessing techniques, like removing stop words, repeating chars, English characters, mentions, punctuation marks, and Arabic diacritics. Also, text normalization has been added.

- English Characters Removal: during this step, all English characters in both lower and upper cases (A-Z, a-z) are removed.
- Stop words Removal: Removing all stop words that may not influence the tweet's meaning.
- Arabic Normalization: Returning chars to their original.
- Arabic Diacritics Removal: Removing all diacritics like [Fatha,Tanwin Fath, Damma, Tanwin Damm, Kasra,Tanwin Kasr,Sukun] .
- Mentions Removal: Removing any mentions like @ from the tweet.
- Repeated Chars Removal: Any repeating characters are removed during this step.
- Punctuations Removal: Removing all punctuation marks like: ". , ' + ! — . . . " " —

#### c) Features Extraction

Feature extraction techniques aim to represent the text's emotional value which is able to help classify the emotions into the right category. Feature extraction is essential before EC from the documents, which can be found in method such as Term Frequency-Inverse Document Frequency (TF-IDF). The next section describes the feature extraction method utilized in our proposed approach:

- TF-IDF

TF-IDF is one in every of the foremost used text feature extraction techniques because it provides helpful insight into the essential features of text documents. during this paper, TF IDF is chosen because the feature extraction technique. It computes the merchandise of the 2 statistics: TF-IDF describes how the word is important to a tweet in an exceedingly collection of tweets. the worth of TF-IDF increases correspondingly to the quantity of times a word appears within the tweets. The more a term occurs in tweets belonging to some

category, the more it's relative there to category. TF-IDF's function is more developed and offers ideal outcomes as it can identify an emotional Arabic term. Figure (5) highlights the characters' number in tweets.

#### d) Experiment

The experiment describes the approach to predicting users' emotions from their tweets. To categorize the tweet into (anger, joy, sadness, and fear), we apply different machine-learning approaches: K-Nearest Neighbor, Decision Tree, Support Vector Machine, Naive Bayes, and Multinomial Naive Bayes. This work has been implemented on a cloud-based environment, "Google Colab," owned by Google. The experiment's first and most essential phase is preprocessing the tweets for training and test sets. mostly, Arabic text needs more preprocessing because of its nature and structure. Therefore, the preprocessing techniques for every tweet are performed for the training and testing phases. We used the dataset of the Arabic tweets presented by Semi-Eval 2018. Then classified, each tweet was placed into one in all four categorizations, given an emotion and a tweet. This dataset includes (934) tweets for the provided emotions: Fear, anger, sadness, and joy. The TF-IDF is extracted from the text and classified using KNN, SVM, DT, NB, and Multinomial NB. We randomly split our dataset into testing and training with 20-80 ratios. the proportion of every class within the dataset is shown in Fig.6. We used the training datasets to point the classifiers. In contrast, (unseen to the model), the test dataset was reserved for examining the structured model to identify the suitability of the trained model. After splitting our dataset into the testing and training process, 747 samples are within the training dataset and 187 within the testing dataset. The results of 5 machine learning models are compared within the result section.

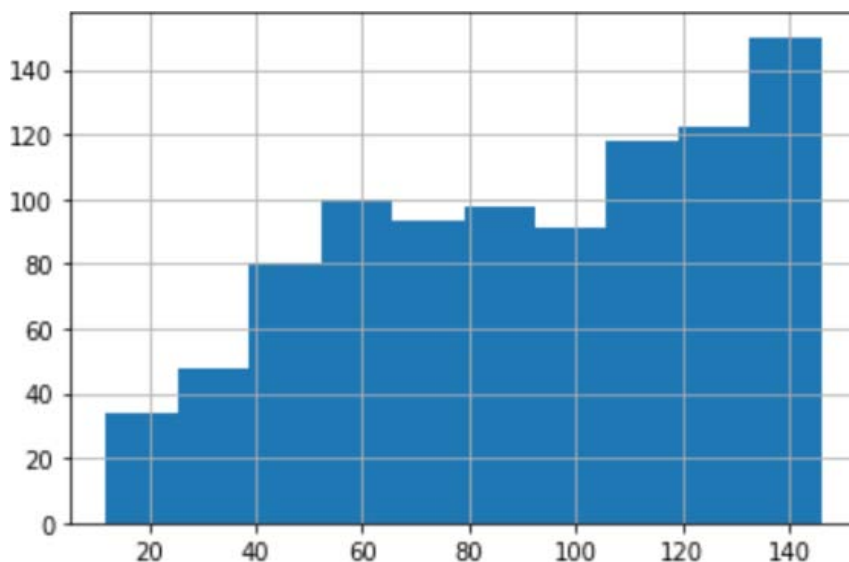


Fig. 5: The number of characters in tweet

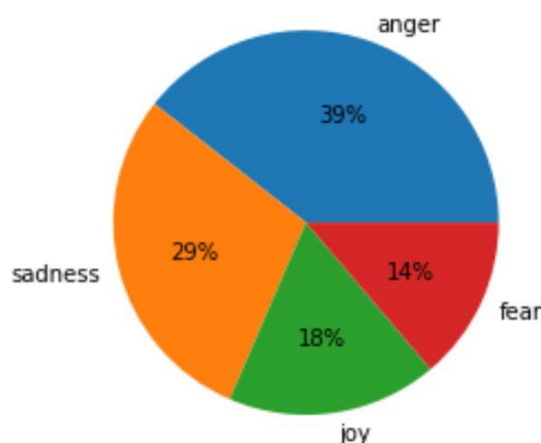


Fig. 6: The percentage of each class in the dataset

Accuracy	$\frac{\text{True Positive} + \text{True negative}}{\text{True Positive} + \text{True negative} + \text{False Positive} + \text{False negative}}$
Recall	$\frac{\text{True Positive}}{\text{True Positive} + \text{False negative}}$
Precision	$\frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$
F-measure	$2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$

Fig. 7: Evaluation Metrics

## V. EVALUATION METRICS

We use precision, accuracy recall, and F- score accuracy, in this study, to measure the EC's performance. Precision, also named positive predictive value, is the documents' number labeled correctly as belonging to the positive class. Sensitivity, or recall, is the documents' number that is not labeled as belonging to the positive class. Another measurement that combines recall and precision is F-score. The F-score indicates how accurate the classifier is (how many instances are correctly classified) and its robustness (it does not miss many instances). The last measure is accuracy, which indicates the suitability of a given classifier[13]. Calculations of the measurements are given in Fig. 7.

## VI. RESULTS AND DISCUSSION

This section reports the performance results and discusses the model. We trained the model on the training dataset for the SemEval- 2018 dataset, and reported on the model performance on the test dataset. The Decision Tree and K- Nearest Neighbor classifier's accuracy was 0.74, the NB and Multinomial NB classifiers obtained 0.69, and the SVM 0.63. The results could be more encouraging regarding accuracy. Fig. 8 displays the accuracy comparison of various machine learning models.

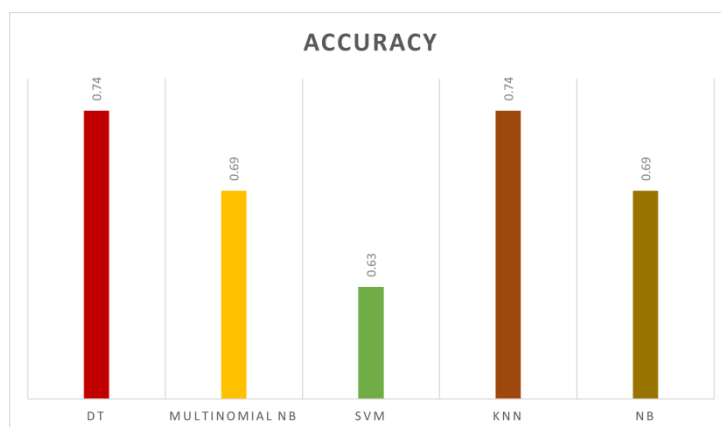


Fig. 8: Results of the accuracy using different Classifiers



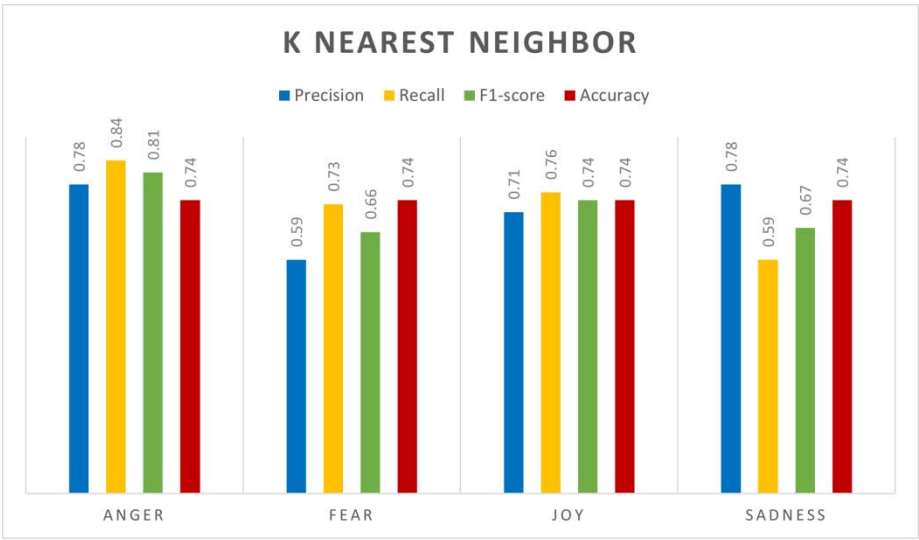


Fig. 9: The findings of the KNN algorithm

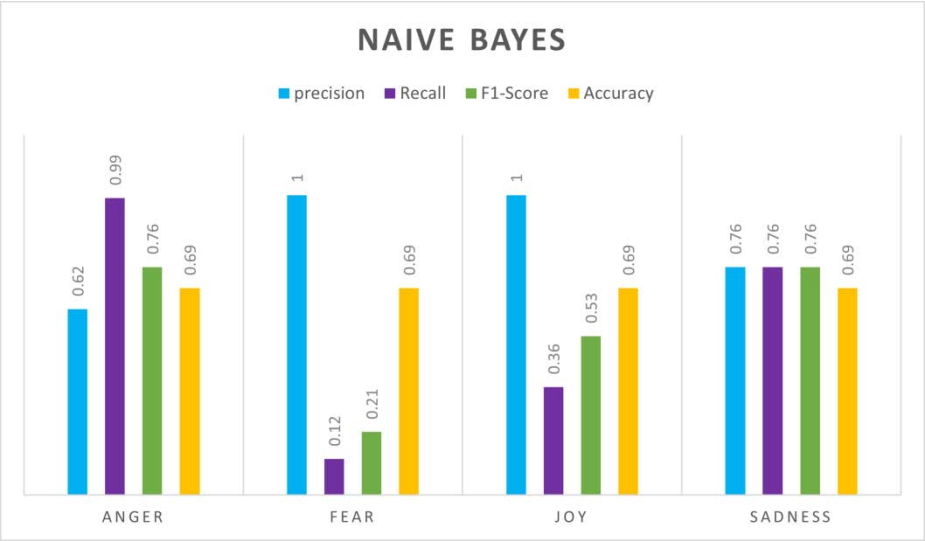


Fig. 10: The NB algorithm's results

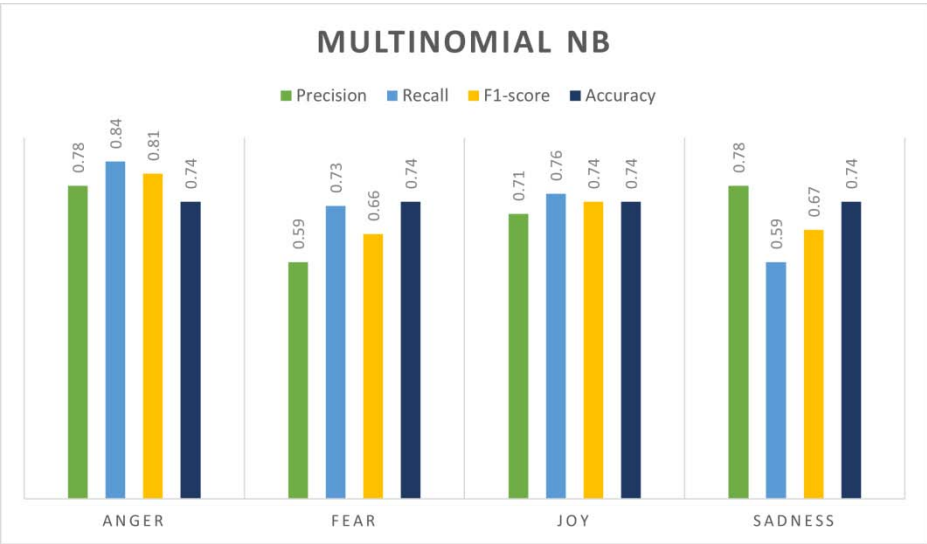


Fig. 11: The Multinomial NB algorithm's results

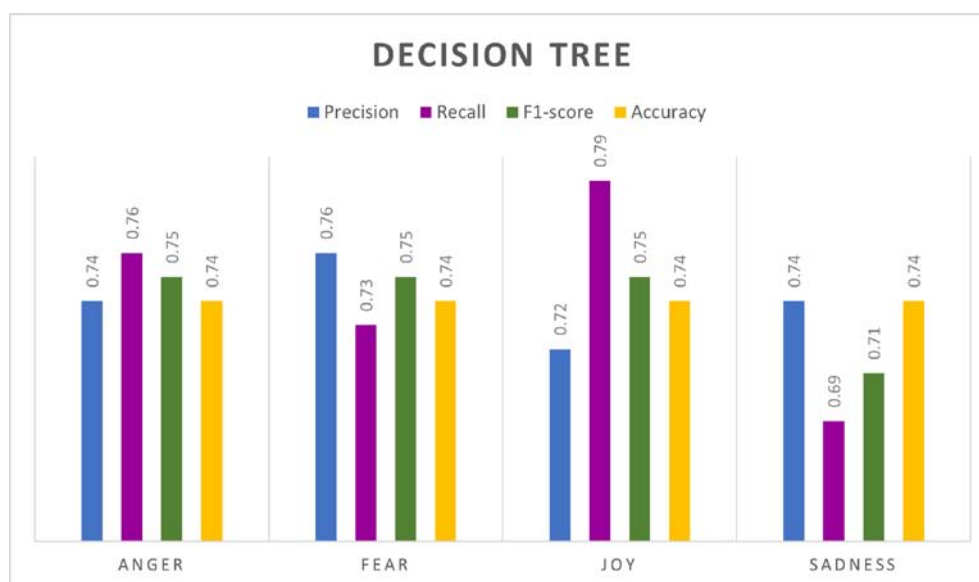


Fig. 12: The DT algorithm's results

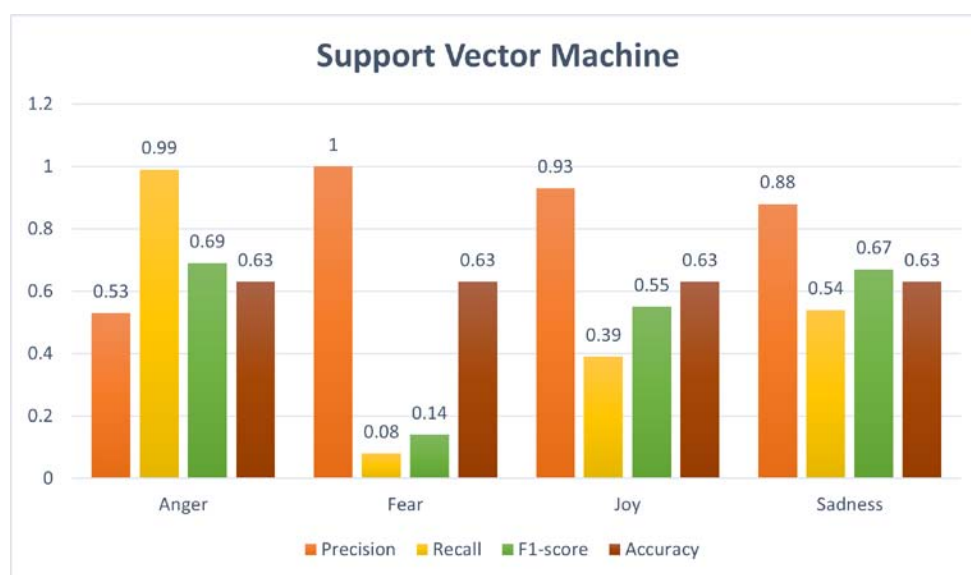


Fig. 13: SVM algorithm's results

## VII. CONCLUSION

EC is a text categorization approach that aims to identify human feelings conveyed through texts. In recent years, Arab users have expressed their emotions on many of the issues raised through the Twitter platform. Therefore, this paper focused on the common classification algorithms such as DT, KNN, SVM, NB, and Multinomial (NB) and applied them to a tweet's dataset as short text content. This study presented the approach for categorizing the emotions of tweets written in Arabic while utilizing the machine learning model. We used the dataset of Arabic tweets presented by SemEval-2018 for El-oc task. This process used (4) emotion categories: Anger, joy, fear, and sadness. The approach achieved acceptable results with 0.74 for each of the KNN and DT, while the NB and Multinomial

NB acquired 0.69; finally, the SVM achieved 0.63. Therefore, future research, including deep learning, is promising, primarily if provided with a large, good, annotated dataset. Also, future research on constructing and finding an Arabic dataset that is labeled correctly will aid and increase the advancements in textual emotion detection, because it will offer a dataset that can be utilized to compare various suggested investigations.

## VIII. FUTURE WORK

The authors intend to examine this model on bigger datasets, assess the outcomes, and employ a hybrid approach that depends on deep learning and machine learning to classify emotions in texts written in Arabic and compare multiple methodologies.

## REFERENCES RÉFÉRENCES REFERENCIAS

1. Mansy, A., Rady, S., Gharib, T. (2022). An Ensemble Deep Learning Approach for Emotion Detection in Arabic Tweets. *International Journal of Advanced Computer Science and Applications*, 13(4).
2. Khalil, E. A. H., El Houby, E. M., Mohamed, H. K. (2021). Deep learning for emotion analysis in Arabic tweets. *Journal of Big Data*, 8(1), 1-15.
3. Elfaik, H. (2021). Combining Context-Aware Embeddings and an Attentional Deep Learning Model for Arabic Affect Analysis on Twitter. *IEEE Access*, 9, 111214-111230.
4. Alswaidan, N., Menai, M. E. B. (2020). Hybrid feature model for emotion recognition in Arabic text. *IEEE Access*, 8, 37843-37854.
5. AlZoubi, O., Tawalbeh, S. K., Mohammad, A. S. (2020). Affect detection from arabic tweets using ensemble and deep learning techniques. *Journal of King Saud University-Computer and Information Sciences*.
6. Hussein, A., Al Kafri, M., Abonamah, A. A., Tariq, M. U. (2020). Mood Detection Based on Arabic Text Documents using Machine Learning Methods. *International Journal*, 9(4).
7. Saad M. M., Jamil, N., Hamzah, R. (2018). Evaluation of support vector machine and decision tree for emotion recognition of malay folklores. *Bulletin of Electrical Engineering and Informatics*, 7(3), 479-486.
8. Hegazi, M. O., Al-Dossari, Y., Al-Yahy, A., Al-Sumari, A., Hilal, A. (2021). Preprocessing Arabic text on social media. *Heliyon*, 7(2), e06191.
9. Nandwani, P., Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, 11(1), 1-19.
10. Boudad, N., Faizi, R., Thami, R. O. H., Chiheb, R. (2018). Sentiment analysis in Arabic: A review of the literature. *Ain Shams Engineering Journal*, 9(4), 2479-2490.
11. Singh, G., Kumar, B., Gaur, L., Tyagi, A. (2019, April). Comparison between multinomial and Bernoulli naïve Bayes for text classification. In *2019 International Conference on Automation, Computational and Technology Management (ICACTM)* (pp. 593-596). IEEE.
12. Rahat, A. M., Kahir, A., Masum, K. M. (2019, November). Comparison of Naive Bayes and SVM Algorithm based on sentiment analysis using review dataset. In *2019 8th International Conference System Modeling and Advancement in Research Trends (SMART)* (pp. 266-270). IEEE.
13. Saad, M. M., Jamil, N., Hamzah, R. (2018). Evaluation of support vector machine and decision tree for emotion recognition of malay folklores. *Bulletin of Electrical Engineering and Informatics*, 7(3), 479-486.