# Review On Using Machine Learning and Deep Learning Algorithms for Emotion Analysis

# مراجعة حول استخدام التعلم الآلي وخوارزميات التعلم العميق لتحليل المشاعر


فتح احمد قائد احمد الجرادي

Fath Ahmed Qayed Aljradi[1]

محمد البارد

Mohammed Albared [2]

عبدالله سعيد غريب

Abdullah Saeed Ghareb [3]



https://doi.org/10.54582/TSJ.2.2.93



(1)Email: Fateh736481212@gmail.com

(2) Faculty of IT&CS University of Saba Region, Marib, Yemen
Email: dr.albared@gmail.com

(3) Faculty of IT&CS University of Saba Region, Marib, Yemen
Email: aghurieb@usr.ac

**الملخص:**

في الوقت الحاضر، أصبحت مشاركة اللحظات على الشبكات الاجتماعية أمرًا واسع الانتشار. ويتم مشاركة الأفكار والخواطر والذكريات الدقيقة للتعبير عن مشاعرنا من خلال النص دون استخدام الكثير من الكلمات. ونتيجة لذلك، أصبح تحليل البيانات النصية لوسائل التواصل الاجتماعي ذا أهمية متزايدة، لأنه يحتوي على أحدث المعلومات حول ما يفكر فيه الناس. على سبيل المثال، يعد تويتر مصدرًا غنيًا للبيانات التي يمكن للمؤسسات استخدامها لتحليل آراء الأشخاص ومشاعرهم وعواطفهم. وعادةً ما يوفر تحليل المشاعر صورة أكثر شمولاً لمشاعر المؤلف. كما تهتم المنظمات والأفراد باستخدام وسائل التواصل الاجتماعي لتحليل آراء الناس واستخلاص المشاعر والعواطف، مما يؤدي بالتالي إلى معرفة توجهات الناس حول موضوع معين. ولا يحظى الكشف عن المشاعر إلا باهتمام قليل للغاية، والقليل جدًا من الأبحاث حتى الآن اختبرت فئة المشاعر في النص، وخاصة المحتوى المكتوب باللغة العربية، وأيضا تؤثر البيانات غير المتوازنة التي تحتوي على نصوص عربية على أداء عملية التصنيف. ولذلك، فقد اكتسب تحليل المشاعر القائم على النص الكثير من الاهتمام في الآونة الأخيرة. تقدم هذه الورقة مراجعة منهجية للأدبيات الموجودة في تحليل العاطفة القائم على النص . (**TBEA**) للإجابة على القضايا البحثية الرئيسية، نظرت هذه الدراسة بعناية في أكثر من ٦٠ منشورًا بحثيًا. بالإضافة إلى ذلك، فهو يتناول منهجيات تحليل المشاعر **EA** المستخدمة في دراسات مختلفة.

**الكلمات المفتاحية:** تحليل الانفعالات، التعلم العميق، التعلم الآلي، اللغة العربية

## Abstract

Nowadays, sharing moments on social networks has turned out to be something widespread. Sharing ideas, thoughts, and precise recollections to explicit our feelings through text without the use of loads of words. As a result, social media text data analysis is becoming increasingly important, as it contains the most up-to-date information on what people are thinking. For example, Twitter is a rich source of data that organizations can use to analyze people's opinions, sentiments, and emotions. Emotion analysis usually provides a more comprehensive picture of an author's feelings. Organizations and individuals are also interested in using social media to analyze people's opinions and extract feelings and emotions, which thus leads to knowing people's orientation on a specific topic. Emotion detection gets extraordinarily little attention. Very little research so far has tested the class of feelings in text, especially Arabic written content. unbalanced data that contains Arabic texts affects the performance of the classification process. Therefore, text-based emotion Analysis has gained a lot of attention in recent times. The paper presents a systematic literature review of the existing literature in Text-Based Emotion Analysis (TBEA). To answer the main research issues, this study has carefully looked at over 60 research publications. Additionally, it goes over the numerous TBEA methodologies used in different study disciplines. A summary of several emotion models and methods.

**Keywords**: Emotion analysis, Deep learning, Machine learning, Arabic language

## 1. Introduction

Emotion Analysis (EA) is a subtask of Natural Language Processing (NLP) that aims to analyze huge data to detect people's opinions and emotions. Emotion analysis, or the detection of more complex feelings, is a relatively new field that presents new challenges in addition to those faced by sentiment analysis where Emotion analysis is a more fine-grained classification [1-3]. Existing Arabic datasets for multi-label text-based emotion analysis suffer from a high level of class imbalance [3] where the number of instances of some given class is very high, while in the other classes, the number of instances is low. In real life, the distribution of examples (training tweets) is skewed since comments that belong to some of the emotion classes appear infrequently. This poses a difficulty to learning algorithms, as they are biased towards the majority classes. Yet most of that text-based Arabic emotion analysis work assumes balanced sample sizes for each emotion class, which is not consistent with reality [3-5]. Supervised learning is not supported effectively in Arabic emotion analysis as datasets with multi-label emotion labels are too small and too imbalanced. Thus, supervised learning methods which are suitable for balanced classification, fail to achieve their intended effects with unbalanced data and directly affect the performance of overall emotion analysis. With not many works focused on imbalanced class problem distribution on emotion analysis in general [6-9] and the absence of any thorough study on the effect of imbalanced classes in Arabic emotion analysis, this work addresses the class imbalance problem, one of the most challenging issues in Arabic multi-label text-based emotion analysis. In addition, Emotional analysis of Arabic is still in its infancy; indeed, many dialects are not covered by researchers and few emotional resources exist which does not encourage research in the field to balance work done in other languages such as English. The objectives of this paper are to present a systematic literature review of the existing literature in TBEA. Additionally, it goes over the numerous TBEA methodologies used in different study disciplines. A summary of several emotion models and methods. The remaining sections of this paper are organized as follows; Section (2) highlights some theoretical Background on emotion analysis Section (3) presents a literature review. Section (4) presents Machine learning and deep learning algorithms used in emotion analysis, Finally, the conclusion.

## 2. Theoretical Background

With the rapid growth of web applications, such as E-commerce platforms and substantial social media comments in various fields, an urgent need to deal with this massive amount of web data and automatically extract helpful information has arisen. Sentiment analysis models play a significant role in this task. Sentiment analysis is a computational field within natural language processing (NLP) concerned with people's sentiments and opinions toward objects such as services, persons, products, events, and organizations [5]. Therefore, the analysis of this user-generated content provides valuable information about the opinion of users about a large variety of topics and products, allowing firms to address typical marketing problems as, for instance, the evaluation of customer satisfaction or the measurement of the impact of a new marketing campaign on brand perception. Moreover, the analysis of customers' opinions about a certain product can be a driver for open innovation, as it helps business owners find out possible issues and suggest new interesting features [10].

The explosive growth of social networking and commenting positively or negatively affects critical business processes such as customer support and satisfaction, brand management, and the reputation of individuals or organizations, as well as affecting product design and marketing. Therefore, sentiment analysis is essential in the industry and business. Feedback is useful for companies to know the satisfaction of users with the service they provide. It is also important for customers to know what people have said about the disadvantages and advantages of this service before purchasing it, and for institutions to provide better services there is a need to discover the different feelings expressed by people and then use this information to make recommendations appropriate to specific needs to their clients. Consequently, a large number of individuals are utilizing social media in the twenty-first century at a rapid pace due to its dependability, speed, and ease of use. As a result, a lot of large sectors are utilizing social media to examine client feedback, examine customer sentiment, and comprehend customer behavior [11].

Emotion Analysis (EA) is a subtask of Natural Language Processing (NLP) that aims to analyze huge data to detect people's opinions and emotions. This field has gained growing interest in the public and private sectors that led to the occurrence of many challenges, especially those related to the Arabic language [12]. And emotions are mainly expressed using language; however, some emotions, such as joy, fear, and sadness, are more fundamental than others and can be expressed differently [5]. Emotion analysis presents new challenges in addition to those of sentiment analysis. The issue concerns the representation of emotions, as researchers in the field of psychology still argue about classifying emotions into a range of distinct categories. In contrast to sentiment analysis where polarity is represented using a single scheme (positive, negative and neutral), many sentiment recognition schemes are adopted by the research community. Emotion analysis can be performed in a variety of ways and has a wide range of applications. The three main techniques are lexicon-based, machine-learning-based, and deep learning-based emotion analysis used for emotion analysis. Each has its own set of benefits and drawbacks [13].

## 3. Literature Review

Emotions are reactions that human beings experience in response to events or situations. Emotions are human created states that represent assessments of oneself, one's surroundings, and other social actors [14]. The situation that elicits an emotion determines the kind of feeling the individual feels. For example, happiness is felt when one gets good news, while fear arises when one feels threatened. Our daily lives are significantly impacted by our emotions. Whether we are pleased, angry, depressed, bored, or dissatisfied influences the decisions we make. We also select interests and pastimes according to the feelings they arouse. Being aware of our emotions may make navigating life easier and more stable, and emotions influence perception, behavior, and learning in all three domains through a combination of these and other effects [15].

Emotion analysis, also known as sentiment analysis, is a branch of artificial intelligence and natural language processing (NLP) focused on understanding, detecting, and interpreting emotions expressed in text, speech, images, or other forms of communication. It plays a pivotal role in understanding human sentiment, behavior, and responses across various domains [16]. Emotion analysis has garnered significant attention due to its relevance across diverse domains, including social media, customer feedback analysis, healthcare, and human-computer interaction. Emotion analysis involves the automated extraction of emotional states, sentiments, or attitudes from text, speech, or other data sources. It is crucial in understanding human interactions, decision-making processes, and the overall mood or sentiment prevalent in communication. The significance of emotion analysis spans across industries, including marketing, healthcare, social media, customer service, and beyond. For text-based emotion analysis, several approaches have been proposed rule-based approaches, classical learning-based approaches, deep learning approaches, and hybrid approaches [13].

### 3.1 Emotion Analysis

Emotion analysis, also referred to as sentiment analysis or affective computing, is a burgeoning field within natural language processing and computational linguistics. Emotion plays a fundamental role in human communication, influencing decisions, behaviors, and interactions. Analyzing emotions expressed in textual data has become increasingly important for various applications such as market research, customer feedback analysis, mental health monitoring, and human-computer interaction. Additionally, it extends to analyzing non-textual data like images, audio, or video to discern emotions from facial expressions, voice tone, or gestures [17]. Overall, emotion analysis aims to computationally understand and interpret the emotional content within data, providing valuable insights into human sentiment, behavior, and perception across a wide range of applications and industries, Figure 1 shows types of emotion analysis and Table 1 shows it in details.


https://doi.org/10.54582/TSJ.2.2.93

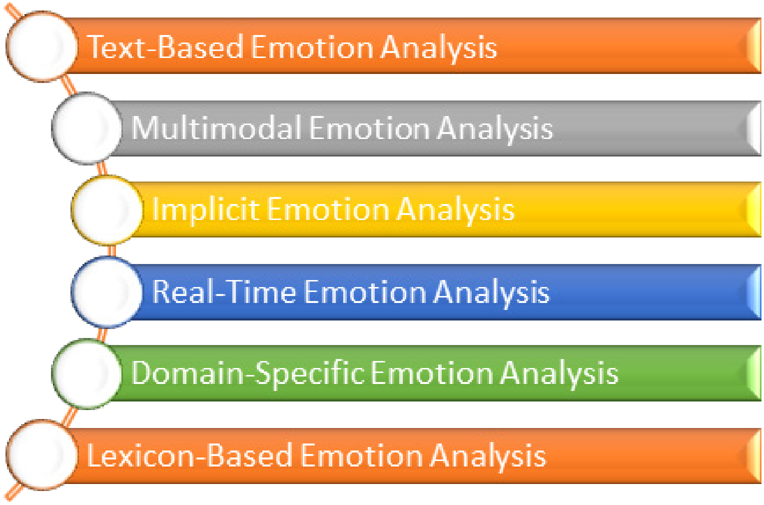


Figure 1 Types of Emotion Analysis

Table 1: Types of Emotion Analysis

| Type of Emotion | About it |
|---|---|
| Text-Based Emotion Analysis | Texts are utilized as a fundamental form of communication to express feelings and ideas, either openly or implicitly. One of the main ways we communicate our thoughts and feelings is via writing. As social media has grown, so too sentimental or emotional states may be expressed through tweets, blog posts, comments, and other social media posts. This led researchers to examine texts from social media platforms and offer a range of techniques for identifying attitudes and emotions. |
| Multimodal Emotion Analysis | Opinions are increasingly being shared online in the form of videos rather than text alone. this has led to SA using multiple modalities, termed multimodal sentiment analysis (MSA) [17]. Too audio-based emotion analysis analyzes emotions conveyed through voice tone, pitch, and speech patterns to discern emotional states in audio data, visual-based emotion analysis examines facial expressions, body language, and gestures to identify emotions portrayed in images or videos [18]. |

| | |
|---|---|
| Implicit Emotion Analysis | Contextual understanding focuses on understanding emotions expressed indirectly or implicitly in text, which may include sarcasm, irony, or implied sentiments not explicitly stated. fine-grained emotion analysis aims to recognize nuanced emotions beyond basic categories, such as subtle emotional states or complex mixtures of emotions [19]. |
| Real-Time Emotion Analysis | Real-time human emotion recognition is a difficult and demanding task that involves recognizing emotions on the face and in speech. Overly, Dynamic Emotion Recognition detects emotional clues in live encounters, video conferences, or streaming content using real-time emotion analysis [20]. |
| Domain-Specific Emotion Analysis | Healthcare and mental health analyze emotions in patient feedback, assessing mental health sentiments, or developing emotion-aware healthcare applications. customer feedback analysis understands customer sentiments towards products/services through reviews, surveys, and feedback analysis. social media monitoring monitors emotions in social media posts to gauge public opinion, sentiment trends, or brand perception. |
| Lexicon-Based Emotion Analysis | Sentiment Lexicons use predefined lexicons or dictionaries containing words associated with specific emotions to assign sentiments or emotions to text based on word sentiment scores [21]. |

Emotion analysis, also known as sentiment analysis, which encompasses various types based on the nature of data and the methodologies used to analyze emotions. Each type of emotion analysis focuses on different aspects of emotional content and utilizes various techniques, methodologies, and data sources to interpret and categorize emotions expressed in textual, verbal, or visual forms of communication [22]. These types collectively contribute to a holistic understanding of emotions conveyed in diverse data modalities and some of the main types of emotion analysis.

### 3.2 Emotion Analysis Method/ Techniques

Deep learning and machine learning are the two classification techniques used by most researchers. Overview of different classifiers used in text-based emotion detection.

This section describes different deep learning and machine learning classifiers used to classify emotion purposes in detail.

### 3.2.1 Machine Learning Classifiers

There are three different categories of learning classifiers supervised, unsupervised, and reinforcement learning. While an unsupervised learning classifier works with unlabeled data, supervised learning classifiers have inputs and intended output labels. Reinforcement learning classifiers are self-supervised, which means they learn by trial and error and are often given a large amount of unlabeled data along with a small sample of annotated data for labelling Machine learning-based Approach. Several categories of emotions are assigned to text using machine learning techniques. supervised and unsupervised machine learning algorithms are two different categories. In the bulk of the research we analyzed, supervised machine learning methods were used. This method often starts with text preparation like tokenization, POS tagging, and lemmatization. Then, only the characteristics with the highest information yield are selected once the text's relevant features have been retrieved. The algorithm is then trained using the labels for the selected traits and emotions. In the end, a trained system is used to predict emotions from unobserved data. The authors examined [35] in this approach. Figure 2 illustrates a Machine learning-based Approach. Methods based on machine learning enable systems to develop and learn on their own as a consequence of their experiences.

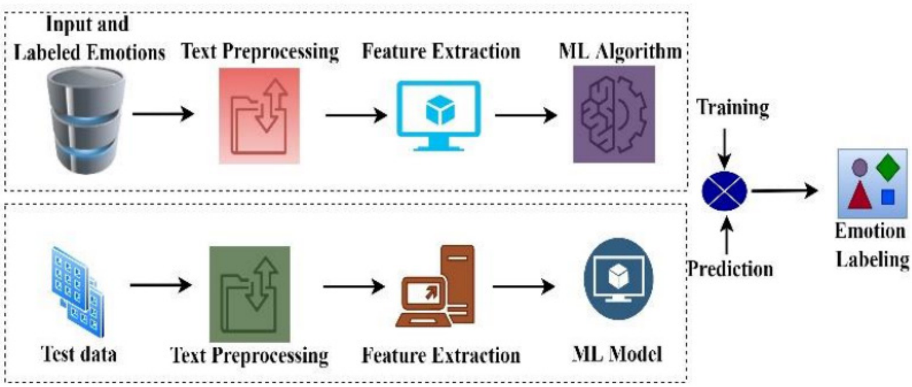


Figure 2 Machine Learning-Based Approach Layout [22]

As shown in Table 2 Machine Learning algorithms. Methods based on machine learning enable systems to develop and learn on their own as a consequence of their experiences. There are three different categories of learning classifiers supervised, unsupervised, and reinforcement learning. Some of the most popular machine learning classifiers are Decision Tree (DT), K Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest (RF), Naive Bayes (NB), Linear Regression (LR), and Multinomial Naive Bayes (MNB).

Table 2: Machine Learning Algorithms

| Algorithm | Discerption |
|---|---|
| Decision Trees (DT) | A decision tree is a rule-based supervised classifier. It has nodes, branches, and directed edges like a tree. Each node denotes an attribute or feature, each branch denotes the decision rules or tests on attributes, and each leaf node exhibits a class label or result. In [23], they developed a machine-learning model for emotion detection from Arabic textual data on social platforms. They used five different machine learning algorithms, namely Decision Tree (DT), KNearest Neighbor (KNN), Naive Bayes (NB), Multinomial Naive Bayes (NB), and Support Vector Machine (SVM) to classify emotions in Arabic tweets. |
| Random Forest (RF) | Random Forest is a supervised classification algorithm. It was constructed with a decision tree method. A group of decision trees that have been integrated into a single entity is referred to as a "forest" to improve the outcome and generate a more precise, reliable forecast. This method is useful for classification and regression issues since it is clear, adaptable, and easy to apply. They created a model using machine learning classifier techniques. Support Vector Machine (SVM), Random Forest (RF), Logistic Regression (LR), K-nearest Neighbors (KNN), AdaBoost, and Naïve Bayes (NB) are introduced in [24]. |
| (k-Nearest Neighbor) | k-NN is one of the simplest categories. The nearest neighbor of K is the meaning of k-nearest neighbor. The main idea behind the k-NN method is that samples are categorized according to the majority votes of their neighbors, with samples being assigned to the class that is most prevalent among their k-closest neighbors. Based on one or more recent samples, the approach identifies the type of sample that may be categorized. To calculate how far a property was from its neighbors, Euclidean distance was utilized. The KNN method is a slow, non-parametric technique that does not rely on explicit training before classification and makes no assumptions about the distribution of the underlying data. Regression and classification are two uses for it [22]. |

| | |
|---|---|
| Support Vector Machine (SVM) | A supervised learning model is an SVM. It is employed to address classification and regression issues. Despite having the ability to handle non-linear data in a high-dimensional feature space, it is mostly employed to arrange linearly separable data. In [25], they present a technique for figuring out whether the sentiments of the clients are neutral, negative, or neutral. They compare the Nave Bayes classifier (NB) with the Support Vector Machine (SVM) to categorize restaurant patron satisfaction in Jakarta. |
| Naïve Bayes (NB) | One of the simplest classifiers used for text categorization is the Nave Bayes classifier. The Nave Bayes assumption is derived from the Bayesian theory of probability and holds that, given the object's class label, an object's characteristics are conditionally independent techniques in classical machine learning and the ensemble method that used an unsupervised approach and NLP modules to analyze sentiment including NB, LR, SVM, RF, and others introduced in [26]. |

### 3.2.2 Deep Learning Classifiers

Deep learning teaches neural networks how to construct complicated hypotheses out of simpler ones. Preprocessing procedures for the emotion dataset include tokenization, stop word removal, and lemmatization. The embedding's are then created. Tokens are represented in this instance by numbers. These vectors are then given to deep neural network layers with components that are equal to emotion labels via classification, where data forms are found and utilized to estimate labels [27]. Figure 3 illustrates the Deep Learning learning-based Approach.

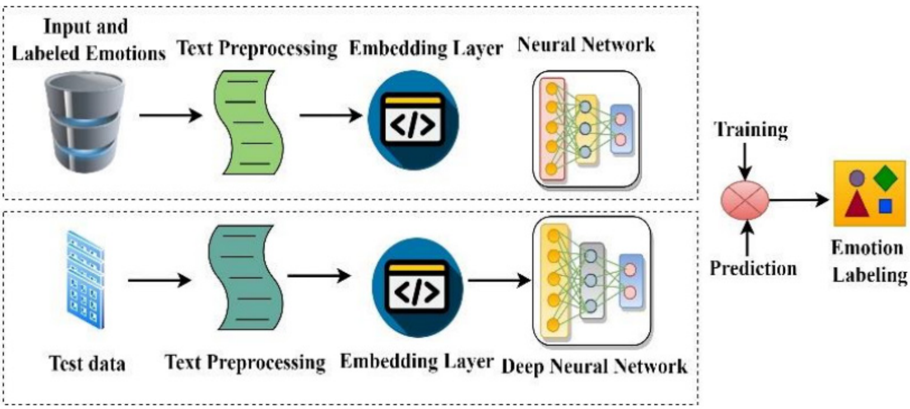


Figure 3: Deep Learning-Based Approach Layout[22].

#### 3.2.2.1 Long Short Term Memory (LSTM)

The LSTM is a type of RNN that solves issues more successfully than the basic RNN by retaining long-term dependency. When it comes to addressing the vanishing gradient problem, LSTM proves to be useful [28, 29]. Like RNN, LSTM also featured a chain-like design, but it used several gates to tightly control how much information could enter each node state. Preprocessing is done on the input to reorganize it for the embedding matrix. The LSTM, the following layer, contains 200 cells. The last layer is fully linked and has 128 text categorization cells. If there are two classes (negative, positive) to forecast, the last layer employs the sigmoid activation function to decrease the 128-height vector to a one-height output vector. One type of LSTM network is a bidirectional LSTM network. Bi-LSTMs consist of two hidden layers. The first hidden layer processes the input sequence forward, while the second hidden layer processes it backwards. The hidden layers are combined in the output layer, which gives it access to each point's past and future context. Both the LSTM and its bidirectional variants have shown to be of great use. They may discover when and how to forget specific facts, as well as when and why not to utilize particular architectural entrances. Faster learning and better performance are benefits of a Bi LSTM network [35].

#### 3.2.2.2 Gate Recurrent Unit (GRU)

GRUs are proposed as a gating mechanism for RNN [30, 31]. The LSTM architecture is simplified in the GRUs architecture. In contrast, a GRU is different from an LSTM in that it has just two gates and no internal memory. A second non-linearity is also not utilized. The Gated Recurrent Unit, a recurrent neural network that aims to address the GRU, GRU is a variation of the LSTM since both are constructed similarly. GRU uses an update and resets approach to get over the vanishing gradient issue. The update gate is used by the model to determine how much historical data from earlier time steps should be included in subsequent stages. The quantity of data that is deleted is decided by the reset gate. The bi-directional GRU variation known as Bi-GRU is a network.

### 3.2.3 Keyword-based Approach

This technique focused on locating keyword occurrences in a given text and comparing them to the annotations registered in the dataset. The first emotion keyword list is created from lexical databases like WordNet or WordNet-Affect. The dataset is then preprocessed. Then, a specified keyword list and text-based emotion terms are matched using keyword analysis. The impact of the emotion term is then assessed. After checking for any possible negation cues and determining the extent of the negation, the emotion tag computation is carried out [32].

### 3.2.4 Rule-based Approach

This approach utilizes linguistic rules to recognize the emotions from the text First, text preparation is applied to the dataset. Data cleansing, tokenization, POS tagging, and other processes are involved. Then, using a statistic and language principles, rules for extracting emotions are constructed. Then, each word is associated with its probabilistic affinity. Later, the most effective rule was used to identify emotion labels in the dataset [22].

### 3.3 RELATED WORKS

In [33], they presented many Machine Learning (ML) algorithms that have been utilized, such as Naive Bayes, Support Vector Machine, and Decision Tree (DT), based on Sentiment analysis the Airline reviews dataset. In [34] they presented Systematic Review Machine-Learning-Based Text Classification. Emotion analysis can be performed in a variety of ways and has a wide range of applications. The three main techniques are lexicon-based, machine-learning-based, and deep learning-based emotion analysis used for Emotion analysis. Each has its own set of benefits and drawbacks. A full survey about emotion analysis research work, used techniques, and available resources are introduced in [35]. A systematic review of applications of natural language processing and future challenges with special emphasis on text-based emotion detection is introduced in [36]. In [37], They utilized decision tree (DT), support vector machine (SVM), artificial neural networks (ANN), K-nearest neighbors (KNN) and Naïve Bayes (NB) besides ensemble models like random forest (RF) and gradient boosting (GB), which use bagging and boosting methods, three sampling approaches based on ensemble hybrid sampling with bagging and boosting machine learning approach for imbalanced data.

In [38], they presented an approach used to classify emotions in Arabic tweets, the model implements a deep Convolutional Neural Network (CNN), and the architecture deep learning approach is an end-to-end network with word, sentence, and document vectorization steps. In [5], they used optimization BiLSTM network for multi-label Arabic emotion analysis and employed a CBOW word embedding model for word representation. A systematic review of applications of natural language processing and future challenges with special emphasis on text-based emotion detection is introduced in [36]. In [39], they address the problem of Arabic affect detection (multi-label emotion classification) by combining the transformer-based model for Arabic language understanding AraBERT and an attention-based LSTM-BiLSTM deep model. AraBERT generates the contextualized embedding, and the attention-based LSTM-BiLSTM determines the label-emotion of tweets. In [10], they investigated the use of Bidirectional Encoder Representations from Transformers (BERT) models for both sentiment analysis and emotion recognition of Twitter data.

In [5], they introduce an optimization BiLSTM network for multi-label text-based Arabic emotion analysis. They used different preprocessing procedures and employed a CBOW word embedding model for word representation. In [40], they introduce a methodology for automatic text annotation to label Arabic text data as multi-label based on the extracted seed key phrase clusters. They reduced the feature dimension by using the Bi-gram alphabet vector representation to construct the document vectors.

In [4], they introduce a model based on three state-of-the-art deep learning models. Two models are special types of Recurrent Neural Networks RNNs (Bi-LSTM and Bi-GRU), and the third model is a pre-trained language model (PLM) based on BERT, and it is called MARBERT transformer.

In [41], the present method consists of three main modules: I Embedding layer-based Word Representation, Bi-LSTM-based Forward and Backward Context Information Saving, and Sigmoid Layer-based Classification, for emotion recognition that takes into account five main emotions Joy, Sadness, Fear, Shame, Guilt. In [42], they enhanced Long Short-Term Memory (ELSTM) For the emotional identification of Twitter data. In [43], they used LSTM, SVM, and nested LSTM to identify multiclass labels for emotion and achieved. In [44], they classified emotions into seven kinds: (fear, anger, love, joy, surprise, thankfulness and sadness) using LSTM AND nested LSTM. In [45], they used the method Naïve Bayes, support vector machines, artificial neural network (ANN), and recurrent neural network (RNN). In [46], they use multi-head attention with bidirectional long short-term memory and convolutional neural network (MHA-BCNN).

There are many uses for emotion analysis, and it may be done in many different ways. For emotion analysis, there are three main techniques: lexicon-based, machine-learning-based, and deep learning-based. Each offers a unique set of advantages and disadvantages. A thorough overview of emotion analysis research projects, methodologies, and resources is provided in [35]. In [47], the researchers applied multiple methods for analyzing emotions in Arabic text, including bidirectional GRU_CNN (BiGRU_CNN), conventional neural networks (CNN), and an XGBoost regressor (XGB). They gathered a dataset of Tweets by utilizing the Twitter API and conducting searches using emotion-related keywords. The findings displayed a Pearson coefficient of 69.2%, showcasing a substantial enhancement of 0.7% compared to the performance of prior best-performing models.

In [38], they presented an approach used to classify emotions in Arabic tweets, the model implements a deep Convolutional Neural Network (CNN), and the architecture deep learning approach is an end-to-end network with word, sentence, and document vectorization steps. The network has seven layers: word vectorization at the input, sentence vectorization at the convolution layer, max pooling, flattening, concatenation, classification at the Dense with ReLu activation layer, and classification at the 4-neuron SoftMax output layer. Additionally, they assessed three machine learning algorithms: multilayer perceptron, Support Vector Machines, and Naïve Bayes. The SemiEval Arabic tweets dataset was utilized by them. There are 1400 tweets in total for each of the four emotions included in the dataset: fear, joy, sorrow, and rage. In [39], they combine an attention-based LSTM-BiLSTM deep model with the transformer-based AraBERT model for Arabic language understanding to address the issue of Arabic affect detection (multi-label emotion categorization). The label-emotion of tweets is determined by the attention-based LSTM-BiLSTM, whereas AraBERT creates the contextualized embedding. Their suggested strategy performs better than the eight baseline techniques. It obtains a noteworthy accuracy rate of 53.82% on the SemEval2018-Task1.In [10], They looked at sentiment analysis and emotion recognition using Bidirectional Encoder Representations from Transformers (BERT) models using Twitter data. In [5], they provide an optimized BiLSTM network for Arabic emotion identification based on multi-label text. For word representation, they used a CBOW word embedding model and other preprocessing techniques. In [40], they provide an artificial text annotation system that uses the derived seed key phrase clusters to label Arabic text data as multi-label. They created the document vectors by reducing the feature dimension through the usage of the Bi-gram alphabet vector representation.

In [4], they introduce a model based on three state-of-the-art deep learning models. Two models are special types of Recurrent Neural Networks RNNs (Bi-LSTM and Bi-GRU), and the third model is a pre-trained language model (PLM) based on BERT, and it is called MARBERT transformer. In [41], the present method consists of three main modules: I Embedding Layer-based Word Representation, Bi-LSTM-based Forward and Backward Context Information Saving, and Sigmoid Layer-based Classification, for emotion recognition that takes into account five main emotions Joy, Sadness, Fear, Shame, Guilt. In [45], They used the method Naïve Bayes, support vector machines, artificial neural network (ANN), and recurrent neural network (RNN). Figure 4 shows the methodology of Arabic text-based emotion detection and analysis and steeps it.

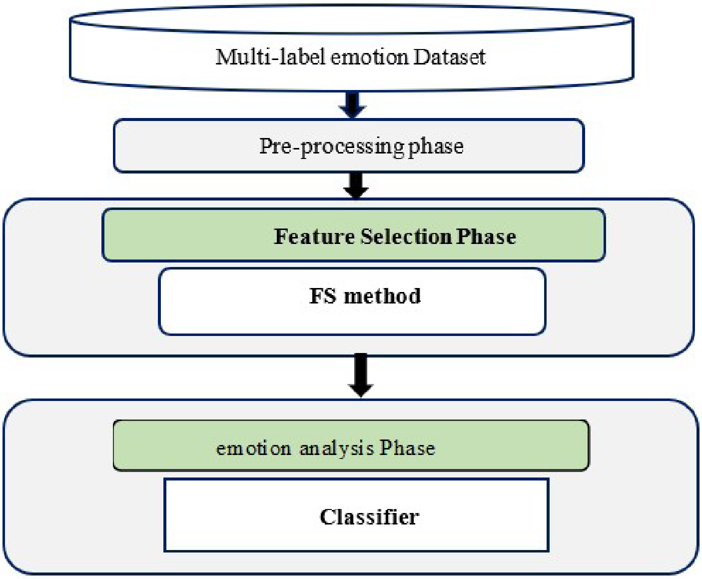


Figure 4: methodology Arabic text-based emotion detection and analysis

## 4. Machine Learning and Deep Learning Algorithm Used in Emotion Analysis

Machine learning classifiers are used extensively and comprehensively in the field of text-based classification. Some of the most popular machine learning classifiers are Naïve Bayes (NB), Support Vector Machine (SVM), Decision Tree (DT), artificial neural networks (ANN), and K-nearest neighbors (KNN) [29]. Machine learning uses computing techniques to extract knowledge directly from a huge number of training examples or samples. As more examples are provided for learning, these algorithms adaptively find the best answers and typically become more efficient. Deep learning classifiers improve accuracy and performance by automatically learning and extracting information. Some of the most popular deep learning classifiers are Convolutional Neural Networks, Recurrent Neural Networks, BERT, Bi-LSTM, GRU, and pre-trained models. A strategy based on deep learning allows for unsupervised learning from unlabeled or unstructured data by using layers of neural networks. It is a machine learning subclass in artificial intelligence. As shown in Table 3, some studies used machine learning and deep learning algorithms to analyze emotions, and a set of algorithms were used to analyze data and identify emotions present in texts, with or without censorship depending on the type of algorithm used in the research.


https://doi.org/10.54582/TSJ.2.2.93

Table 3: Studies that used machine learning and deep learning algorithms for emotion analysis

| Author | Methodology | Algorithm used |
|---|---|---|
| Patel, A., P. Oza [33] | Machine learning | Naive Bayes, Support Vector Machine, and Decision Tree (DT) |
| Malek, N.H.A., et al. [37] | Machine learning | decision tree (DT), support vector machine(SVM), artificial neural networks (ANN), K-nearest neighbors (KNN) and Naïve Bayes (NB) |
| Baali, M. and N. Ghneim [38] | deep learning | Convolutional Neural Networks (CNN) |
| Khalil, E.A.H., E.M. El Houby, and H.K. Mohamed [5] | deep learning | bidirectional long short-term memory (BiLSTM) |
| Chiorrini, A. , et al. [10] | deep learning | Bidirectional Encoder Representations from Transformers (BERT) |
| Mansy, A., S. Rady, and T. Gharib [4] | deep learning | Bi-LSTM and Bi-GRU) and MARBERT |
| Asghar, M.Z., et al.[41] | deep learning | bidirectional long-term short-term memory (BiLSMT) |
| Karna, M., D.S. Juliet, and R.C. Joy. | deep learning | Long Short-Term Memory(LSTM) |
| Elfaik, H. [39] | deep learning | LSTM-BiLSTM, AraBERT |

## Conclusion

Emotion analysis has become important in recent years to know people's opinions and feelings about a particular topic. This article provides a comprehensive review of the literature on TBEA using Emotion analysis techniques. Existing literature shows different artificial intelligence approaches used like deep learning-based, machine learning-based, rule-based, and keyword-based approaches for TBEA. Deep learning and machine learning-based approaches are trending with the help of available datasets and automated feature extraction methods. In this research area, future research study is required to improve development and contribution. Future research should focus on creating emotional databases for various colloquial and Arabic dialects. As well as developing ensemble learning models by integrating advanced deep learning models.

## References


1. Arcan, M. and P. Buitelaar. *MixedEmotions: Social Semantic Emotion Analysis for Innovative Multilingual Big Data Analytics Markets*. in *Proceedings of the 18th Annual Conference of the European Association for Machine Translation*. 2015.

2. Dvoynikova, A., O. Verkholyak, and A. Karpov. *Emotion Recognition and Sentiment Analysis of Extemporaneous Speech Transcriptions in Russian*. 2020. Cham: Springer International Publishing.

3. Alswaidan, N. and M.E.B. Menai, *Hybrid Feature Model for Emotion Recognition in Arabic Text.* IEEE Access, 2020. **8**: p. 37843-37854.

4. Mansy, A., S. Rady, and T. Gharib, *An Ensemble Deep Learning Approach for Emotion Detection in Arabic Tweets.* International Journal of Advanced Computer Science and Applications, 2022. **13**(4).

5. Khalil, E.A.H., E.M. El Houby, and H.K. Mohamed, *Deep learning for emotion analysis in Arabic tweets.* Journal of Big Data, 2021. **8**(1): p. 1-15.

6. Vora, S.V., R.G. Mehta, and S.K. Patel, *Impact of Balancing Techniques for Imbalanced Class Distribution on Twitter Data for Emotion Analysis: A Case Study*, in *Data Preprocessing, Active Learning, and Cost Perceptive Approaches for Resolving Data Imbalance*. 2021, IGI Global. p. 211-231.

7. Jamal, N., et al., *A Deep Learning–based Approach for Emotions Classification in Big Corpus of Imbalanced Tweets.* Transactions on Asian and Low-Resource Language Information Processing, 2021. **20**(3): p. 1-16.

8. Yan, J.L.S. and H.R. Turtle. *Fine-grained Emotion Classification: Class Imbalance Effects on Classifier Performance*. in *2021 International Conference on Computer & Information Sciences (ICCOINS)*. 2021. IEEE.

9. Farsiah, L., Y.-S. Chen, and A. Misbullah. *Multi-Classes Emotion Detection for Unbalanced Indonesian Tweets*. in *2020 International Conference on Electrical Engineering and Informatics (ICELTICs)*. 2020. IEEE.

10. Chiorrini, A., et al. *Emotion and sentiment analysis of tweets using BERT*. in *EDBT/ICDT Workshops*. 2021.

11. Nasir, M.U., U. Rehmat, and I. Ahmad, *Social media analysis of customer emotions in pizza industry.* The Computer Journal, 2023. **66**(7): p. 1777-1783.

12. Nassif, A.B., et al., *Deep learning for Arabic subjective sentiment analysis: Challenges and research opportunities.* Applied Soft Computing, 2021. **98**: p. 106836.

13. Wankhade, M., A.C.S. Rao, and C. Kulkarni, *A survey on sentiment analysis methods, applications, and challenges.* Artificial Intelligence Review, 2022. **55**(7): p. 5731-5780.

14. Peng, S., et al., *A survey on deep learning for textual emotion analysis in social networks.* Digital Communications and Networks, 2022. **8**(5): p. 745-762.

15. Emanuel, A. and E. Eldar, *Emotions as computations.* Neuroscience & Biobehavioral Reviews, 2023. **144**: p. 104977.

16. Zhang, J., et al., *Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review.* Information Fusion, 2020. **59**: p. 103-126.

17. Gandhi, A., et al., *Multimodal sentiment analysis: A systematic review of history, datasets, multimodal fusion methods, applications, challenges and future directions.* Information Fusion, 2023. **91**: p. 424-444.

18. Pan, B., et al., *A review of multimodal emotion recognition from datasets, preprocessing, features, and fusion methods.* Neurocomputing, 2023: p. 126866.

19. Khoshnam, F. and A. Baraani-Dastjerdi, *A dual framework for implicit and explicit emotion recognition: An ensemble of language models and computational linguistics.* Expert Systems with Applications, 2022. **198**: p. 116686.

20. D. Joshi, A.D., A. Khandelwal, S. Kulkarni and S. Mangrulkar, , *"Real Time Emotion Analysis (RTEA)".* International Conference on Artificial Intelligence and Machine Vision (AIMV), , 2021: p. pp. 1-5, .

21. Azam, N., B. Tahir, and M.A. Mehmood, *Sentiment and emotion analysis of text: A survey on approaches and resources.* Language & Technology, 2020. **87**.

22. Kusal, S., et al., *A Review on Text-Based Emotion Detection--Techniques, Applications, Datasets, and Future Directions.* arXiv preprint arXiv:2205.03235, 2022.

23. Aljwari, F., *Emotion Detection in Arabic Text Using Machine Learning Methods.* IJISCS (International Journal of Information System and Computer Science), 2022. **6**(3): p. 175-185.

24. Musleh, D.A., et al., *Twitter Arabic Sentiment Analysis to Detect Depression Using Machine Learning.* Computers, Materials & Continua, 2022. **71**(2).

25. Gunawan, L., et al., *Support vector machine based emotional analysis of restaurant reviews.* Procedia Computer Science, 2023. **216**: p. 479-484.

26. Al Maruf, A., et al., *Emotion detection from text and sentiment analysis of Ukraine Russia war using machine learning technique.* International Journal of Advanced Computer Science and Applications, 2022. **13**(12).

27. Basiri, M.E., et al., *ABCDM: An attention-based bidirectional CNN-RNN deep model for sentiment analysis.* Future Generation Computer Systems, 2021. **115**: p. 279-294.

28. Li, A. and S. Yi, *Emotion analysis model of microblog comment text based on CNN-BiLSTM.* Computational Intelligence and Neuroscience, 2022. **2022**.

29. Sushmitha, M., K. Suresh, and K. Vandana. *To Predict Customer Sentimental behavior by using Enhanced Bi-LSTM Technique*. in *2022 7th International Conference on Communication and Electronics Systems (ICCES)*. 2022. IEEE.

30. Chung, J., et al., *Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling." arXiv, Dec. 11, 2014.* Accessed: May, 2022. **22**.

31. Chung, J., et al., *Empirical evaluation of gated recurrent neural networks on sequence modeling. CoRR. 2014; abs/1412.3555*, 2022.

32. Kusal, S., et al., *AI based emotion detection for textual big data: techniques and contribution.* Big Data and Cognitive Computing, 2021. **5**(3): p. 43.

33. Patel, A., P. Oza, and S. Agrawal, *Sentiment Analysis of Customer Feedback and Reviews for Airline*

*Services using Language Representation Model.* Procedia Computer Science, 2023. **218**: p. 2459-2467.

34. Palanivinayagam, A., C.Z. El-Bayeh, and R. Damaševičius, *Twenty Years of Machine-Learning-Based Text Classification: A Systematic Review.* Algorithms, 2023. **16**(5): p. 236.

35. Uymaz, H.A. and S.K. Metin, *Vector based sentiment and emotion analysis from text: A survey.* Engineering Applications of Artificial Intelligence, 2022. **113**: p. 104922.

36. Kusal, S., et al., *A systematic review of applications of natural language processing and future challenges with special emphasis in text-based emotion detection.* Artificial Intelligence Review, 2023: p. 1-87.

37. Malek, N.H.A., et al., *Comparison of ensemble hybrid sampling with bagging and boosting machine learning approach for imbalanced data.* Indones. J. Elec. Eng. Comput. Sci, 2023. **29**: p. 598-608.

38. Baali, M. and N. Ghneim, *Emotion analysis of Arabic tweets using deep learning approach.* Journal of Big Data, 2019. **6**(1): p. 1-12.

39. Elfaik, H., *Combining Context-Aware Embeddings and an Attentional Deep Learning Model for Arabic Affect Analysis on Twitter.* IEEE Access, 2021. **9**: p. 111214-111230.

40. Elghannam, F., *Multi-Label Annotation and Classification of Arabic Texts Based on Extracted Seed Keyphrases and Bi-Gram Alphabet Feed Forward Neural Networks Model.* ACM Transactions on Interactive Intelligent Systems, 2022.

41. Asghar, M.Z., et al., *A Deep Neural Network Model for the Detection and Classification of Emotions from Textual Content.* Complexity, 2022. **2022**.

42. BABOO, S.S. and M. AMIRTHAPRIYA, *EMOTIONAL ANALYSIS OF TWITTER SOCIAL MEDIA DATA WITH AN EFFICIENT DEEP LEARNING MODEL.* 2022.

43. Karna, M., D.S. Juliet, and R.C. Joy. *Deep learning based text emotion recognition for chatbot applications*. in *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)*. 2020. IEEE.

44. Haryadi, D. and G.P. Kusuma, *Emotion detection in text using nested long short-term memory.* 11480 (IJACSA) International Journal of Advanced Computer Science and Applications, 2019. **10**(6).

45. Mukherjee, P., et al., *Effect of negation in sentences on sentiment analysis and polarity detection.* Procedia Computer Science, 2021. **185**: p. 370-379.

46. Dheeraj, K. and T. Ramakrishnudu, *Negative emotions detection on online mental-health related patients texts using the deep learning with MHA-BCNN model.* Expert Systems with Applications, 2021. **182**: p. 115265.

47. AlZoubi, O., S.K. Tawalbeh, and A.-S. Mohammad, *Affect detection from Arabic tweets using ensemble and deep learning techniques.* Journal of King Saud University-Computer and Information Sciences, 2020.


https://doi.org/10.54582/TSJ.2.2.93