# *Text-Based Emotion Detection: review*

Jannatul Ferdaws Amrin, Asoad Alvi Yanur Saom, Rubaiya Reza Sohana, Sabrina Afrin Toma

Israt Jahan Munny, Shahrin Khan, Rifat Ara

Department of Computer Science & Engineering, Jahangirnagar University

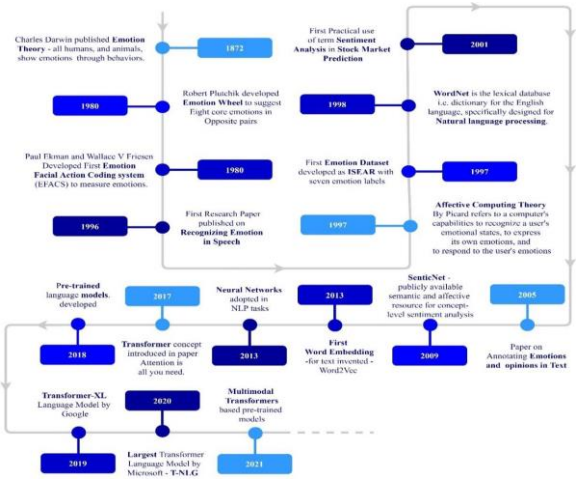
## Introduction

Emotion expressions are the crucial form of communication in interpersonal relationship. It can be expressed into positive emotion, negative emotion or neutral. In general, the positive emotions are expressed as happy, excited, joy, pride and negative emotions such as sad, disgust, fear, depressed and so on. In such a way, the emotions are expressed in various forms to communicate and the rich source of textual information is obtained from the social networking sites such as YouTube, Twitter, and Facebook etc., where people are spending most of their time in posting and expressing their emotions.

Recognizing user's emotions is a major challenge for both humans and machines. Machines need to have accurate ground truth for emotion modeling, and also require advanced machine learning algorithms for developing the emotion models.

Social networking platforms have become an essential means for communicating feelings to the entire world due to rapid expansion in the Internet era. Text communication via Web-based networking media is somewhat overwhelming. Every second, a massive amount of unstructured data is generated on the Internet due to social media platforms. The data must be processed as rapidly as generated to comprehend human psychology, and it can be accomplished using sentiment analysis, which recognizes polarity in texts. It assesses whether the author has a negative, positive, or neutral attitude toward an item, administration, individual, or location. {{00776-6}}

Emotion recognition from text has many applications. Consider for example an employee sending a harsh email to his colleague or superior. A tool that can analyze the email for emotions and alert the employee about its harshness before sending it comes in very handy to protect the employee's state. Consider also an emotion-based search engine that ranks documents according to the emotion requested by the user. Such an engine could prove to be very beneficial to users in a certain emotional state and can improve the effectiveness of the information retrieval process. Other useful tools that can benefit from emotion recognition from text include recommender systems that aim to personalize recommendations based on the user's emotions. {{6486}}.




In addition, many users give feedbacks and reviews various products and services on various e-commerce sites. User's ratings and reviews on multiple platforms encourage vendors and service providers to enhance their current systems, goods, or services. {{00776-6}} Thus it can be be used to extract and classify positive

or negative review of a product in E-Commerce sites like Amazon, Flipkart, Daraz etc. Today almost every industry or company is undergoing some digital transition, resulting in vast amounts of structured and unstructured increase data. The enormous task for companies is to transform unstructured data into meaningful insights that can help them in decision-making (Ahmad et al. 2020). People's active feedback is valuable not only for business marketers to measure customer satisfaction and keep track of the competition but also for consumers who want to learn more about a product or service before buying. {{00776-6}}

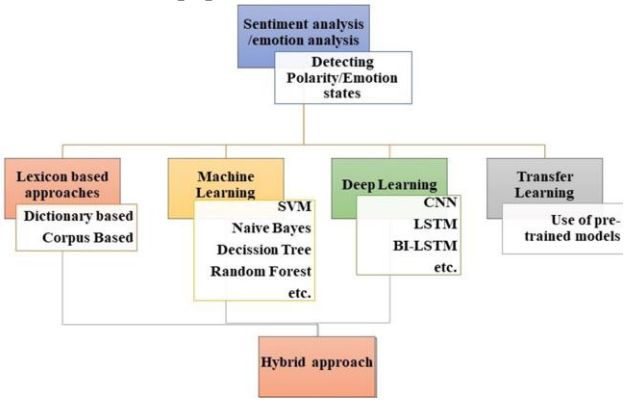
For this reason, we have decided to analyses different research papers which used various existing techniques like Support Vector Machine (SVM), K-Nearest Neighbor (KNN) and different ML approaches for detecting emotion-based systems from text classifier. In this paper, we are presenting a review on the existing emotion detection from text approaches. {{6486}}.

# Methods in Discussion

We have discussed different types of algorithms for our text-based emoticon review paper. Mainly all of the research paper followed these steps to complete the research procedure:

- Collect the data
- Data Cleaning (Emoji Removal and Hashtag Simplifying)
- Tokenization
- Punctuation Marks Removal
- Stop Word Removal
- Vectorization
- Data Splitting
- Apply Models
- Polarity Prediction
- Performance Evaluation
- Final Decision

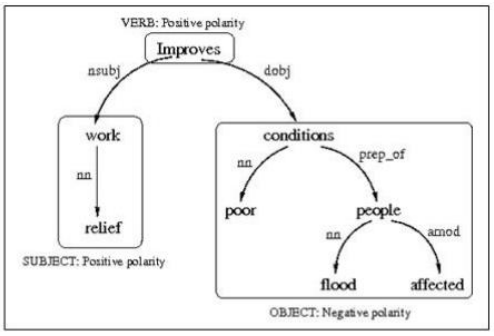
All of the algorithm followed above procedure. We have tried to discuss following algorithms based research paper.




**Support Vector Machine:**

Support Vector Machines (SVM) is a machine learning model proposed by V.

N.Vapnik [10]. The basic idea of SVM is to find an optimal hyperplane to separate two classes with the largest margin from pre-classified data. After this hyperplane is determined, it is used for classifying data into two classes based on which side they are located. By applying appropriate transformations to the data space before computing the separating hyperplane, SVM can be extended to cases where the margin between two classes is non-linear.




The SVM is systematic and motivated by statistical learning theory [6] and Bayesian arguments. The training task involves optimization of a con vex cost function: there are no false local minima to complicate the learning process. The approach has much other benefit, for example, the model constructed has

an explicit dependence on the most informative patterns in the data (the support vectors).

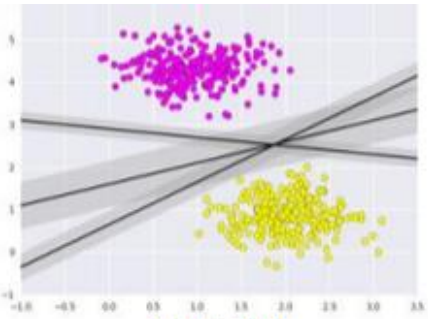


**Fig 1: SVM (a)**

What Support Vector Machines do is consider a region around the line separating two classes, also referred to as a "bounded search space".

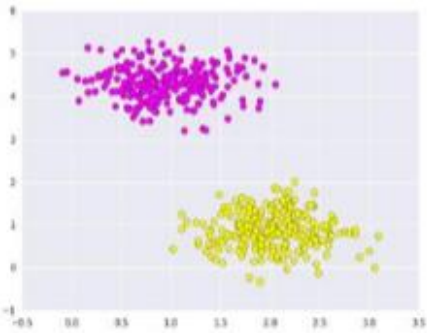


**Fig 2: SVM (b)**

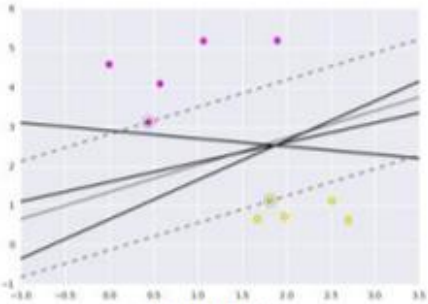


**Fig .3: SVM (c)**

This system recognizes the emotion from the sentence that be inputted from the keyboard. The theory of emotion recognition is the emotion classification with the SVMs. The tool of the SVMs model is the LIBSVM [8]. The system uses the data which has been rightly classed by us for training the classification model and then uses the classification model to predict the sort of emotion, namely negative or non-negative. In this way, the emotion of sentences which be inputted from the keyboard can be recognized.

Support Vector Machine classifier to identify the text orientation, and dealt with the negative sentence via matching negative rules. It also can identify polarity of text. It provides 76.25% accurate result for text-based emotion detection.

## K-Nearest Neighbor (KNN) Algorithm:

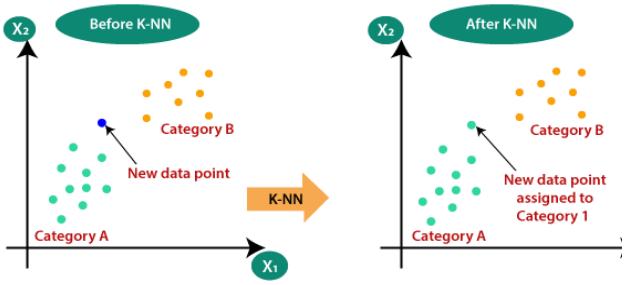
K-Nearest Neighbor is one of the simplest Machine Learning algorithms based on Supervised Learning technique. This algorithm assumes the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available categories.

It stores all the available data and classifies a new data point based on the similarity. This means when new data appears then it can be easily classified into a well suite category by using K-NN algorithm. It can be used for Regression as well as for Classification but mostly it is used for the Classification problems.

K-NN is a **non-parametric algorithm**, which means it does not make any assumption on underlying data. It is also called a **lazy learner algorithm** because it does not learn from the training set immediately instead it stores the dataset and at the time of classification, it performs an action on the dataset.

KNN algorithm at the training phase just stores the dataset and when it gets new data, then it classifies that data into a category that is much similar to the new data.

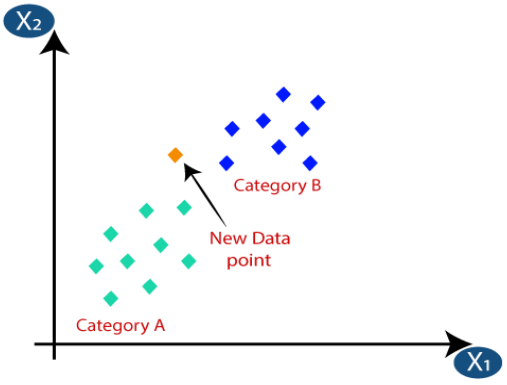
Suppose there are two categories, i.e., Category A and Category B, and we have a new data point x1, so this data point will lie in which of these categories. To solve this type of problem, we need a K-NN algorithm. With the help of K-NN, we can easily identify the category or class of a particular dataset. Consider the below diagram:



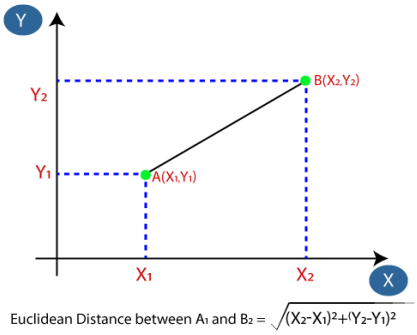

The K-NN working can be explained on the basis of the below algorithm:

- **Step-1:** Select the number K of the neighbor's o
- **Step-2:** Calculate the Euclidean distance of **K number of neighbors**
- **Step-3:** Take the K nearest neighbors as per the calculated Euclidean distance.
- **Step-4:** Among these k neighbors, count the number of the data points in each category.
- **Step-5:** Assign the new data points to that category for which the number of the neighbor is maximum.
- **Step-6:** Our model is ready.

Suppose we have a new data point and we need to put it in the required category. Consider the below image:




- Firstly, we will choose the number of neighbors, so we will choose the k=5.
- Next, we will calculate the **Euclidean distance** between the data points. The Euclidean distance is the distance between two points, which we have already studied in geometry. It can be calculated as:




- By calculating the Euclidean distance, we got the nearest neighbors, as three nearest neighbors in category A and two nearest neighbors in category B. Consider the below image:


- As we can see the 3 nearest neighbors are from category A, hence this new data point must belong to category A.

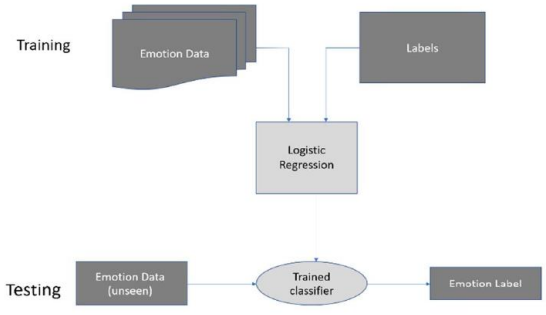
Below are some points to remember while selecting the value of K in the K-NN algorithm: o A very low value for K such as K=1 or K=2, can be noisy and lead to the effects of outliers in the model. o Large values for K are good, but it may find some difficulties.

Steps to implement the KNN Algorithm are given:

- Data Pre-processing step
- Fitting the K-NN algorithm to the Training set
- Predicting the test result
- Test accuracy of the result (Creation of Confusion matrix)
- Visualizing the test set result.

## Logistic Regression:

This paper suggests a supervised machine learning approach based on Logistic Regression (LR) with publicly available emotion dataset, namely ISEAR [4]. The proposed study works on 5 emotion signals and the classifier is trained on the well annotated emotion dataset. The LR algorithm works on training and testing dataset of emotions. Before, making input to the dataset, some preprocessing steps are applied on the input text. There has been used different Python libraries for the implantation of the system. Performance evaluation results are encouraging and shows the effectiveness of the proposed system.




The Logistic Regression (LR) classifier is applied to classify emotions from the input text. Diagrammatic representation is shown that LR performs emotion classification using supervised machine learning approach. Text-based emotions are detected and classified using Logistic Regression (LR) performs by taking training and testing data.

## ML Approach:

ML based classification is called as supervised learning because this process is guided by the labeled training set. There exist various machine learning approaches. The ML approach solves the ED problem by classifying texts into various emotion categories through the implementation of ML algorithms.

**Proposed ML approach: Novel intervention aspect-based (NIAB) analysis**

It classifies the emotion from the twitter dataset. This supervised ML algo uses an activation function to identify output of each node accurately. They are determined through mapping the dependent variable. For a clear direction of slope, a non-linear activation function is introduced called an Exponential Linear Unit (ELU) that has the highest scale with positive value. The exact range between the input and output value is 0.1 and 0.3 which has better output compared to Rectified linear unit (RELU).

**Proposed ML approach: A novel fuzzy clustering model.**

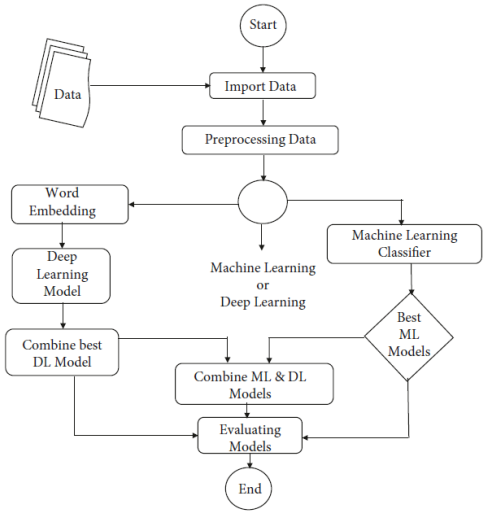
It uses a real-life dataset. A comparative analysis is made with the existing partitioning clustering techniques namely K-Means and Expectation Maximization algorithms based on metrics namely accuracy, precision, recall and execution time.

## Deep Learning Approach:

To identify emotions from text, several methods have been proposed in the past using natural language processing (NLP) techniques: the keyword approach, the lexicon-based approach, and the machine learning approach. However, there were some limitations with keyword- and

lexicon-based approaches as they focus on semantic relations. In this article, they have proposed a hybrid (machine learning + deep learning) model to identify emotions in text.

Convolutional neural network (CNN) and Bi-GRU were exploited as deep learning techniques. Support vector machine is used as a machine learning approach. The performance of the proposed approach is evaluated using a combination of three different types of datasets, namely, sentences, tweets, and dialogs, and it attains an accuracy of 80.11%.




In ML, the preprocessed data will be given as the input to ML classifiers and will show the results of all ML classifiers. Furthermore, it will select the best ML model that gives the highest accuracy. In DL, the data is converted into vector form and given as an input to the DL models. Before that, they have used a trained word vector to make the word embedding matrix and add the embedded layer to the DL model [18]. After performing on individual models, they combine the two best DL models based on accuracy and F1 score.

This research used the prebuilt models of ML and DL. For ML, classifiers like DT, SVM, NB, and RF were built to predict emotions, and for deep learning, we deployed Gated Recurrent Unit (GRU), Bidirectional Gated Recurrent Unit (BiGRU), and Convolutional Neural Network (CNN) to predict emotions.

According to the ML classifier, SVM gives the highest accuracy of 78.97%. In the DL method, the Bi-GRU model achieves the highest accuracy of 79.46%, and the CNN model achieves the highest F1-score of 80.76. The hybrid model has achieved a precision of 82.39, a recall of 80.40, an F1 score of 81.27, and an accuracy of 80.11%.

# Interpretation and Analysis

SVM sometimes overlaps positive and negative emotion when the number of properties for each data point outstrips the number of training data specimens

KNN is simple to implement, robust to the noisy training data and can be more effective if the training data is large. But it always needs to determine the value of K which may be complex some time. The computation cost is high because of calculating the distance between the data points for all the training samples.

Logistic Regression performed better than other similar methods (SVC, KNN, XGBoost) for emotion detection from text. The main drawback of the system is that it needs further exploration with deep learning models, as such technique are now becoming popular in data science community due their automatic feature engineering.

Supervised approach gives more accurate detection than unsupervised Algorithms. Unsupervised is not robust and do not explicitly extract information effectively.

In the deep learning approach, we may get the combination of CNN, Bi-GRU, and LSTM to improve the results. Additionally, we can work on the structure of text sentences and some of the regional languages. Moreover, in this digital world, people's usage of sending text messages,

uploading tweets, and writing online reviews of products have been in great use and demand. Therefore, by having a lot of data, we can make a real-time text-based emotion recognition model to find the emotions or moods of the people. The main drawback of the system is that it needs further exploration with many algorithms and hard to decide.

# Conclusion

Text based emotion plays a significant role in business decision making. Many of the organization and enterprises will take their business decision only based on their customer review. In this study, the overview of different text emotion recognition methods is discussed for extracting text features from text sample, various classifier algorithms are explained briefly.

Deep Learning and SVM is more suitable for this type of text classifiers and lexicons. English text Emotion Recognition has a promising future and its accuracy depends upon the combination of features extracted, type of classification algorithm used and the correct of emotional text database. This study aims to provide a simple guide to the researcher for those carried out their research study in the text-based emotion detection systems.

# Reference


1. M. Shorif Uddin ,Md. Sabab Zulfiker, Nasrin Kabir, Hafsa Moontari Ali, M. Reduanul Haque, Morium Akter and, "Sentiment Analysis Based on Users' Emotional Reactions about Ride-sharing Services on Facebook and Twitter,"

   Proc. of International Joint Conference on Computational Intelligence (IJCCI 2019), 25-26 October 2019, Dhaka, Bangladesh, (Algorithms for Intelligent

   Systems, Springer), pp. 397-408, https://doi.org/10.1007/978-981-15-36076_32

2. Md. Asadur Rahman, Anika Anjum, Md. Mahmudul Haque Milu, Farzana Khanam, Mohammad Shorif Uddin, Md. Nurunnabi Mollah, "Emotion recognition from EEG-based relative power spectral topography using convolutional neural network," Array, (Elsevier), 11, 100072, June 2021. https://doi.org/10.1016/j.array.2021.100072

3. 00776-6 => A review on sentiment analysis and emotion detection from text: Pansy Nandwani1 · Rupali Verma1 https://link.springer.com/article/10.1007/s13278-021-00776-6

4. {{6486}}=> A Review on the Emotion Detection from Text using Machine Learning Techniques: Ubeeka Jain and Amandeep Sandhu https://inpressco.com/a-review-on-the-emotion-detection-from-text-usingmachine-learning-techniques/

5. B. Plaban Kumar Bhowmick and P. Mitra. Classifying emotion in news sentences: Whenmachine classification meets human classification. International Journal on Computer Science andEngineering, 2(1):98–108, 2010 https://www.academia.edu/2642268/CLASSIFICATION_OF_EMOTIONS_FROM_TEXT_USING_SVM_BASED

6. Z. Teng, F. Ren, and S. Kuroiwa, "Recognition of Emotion with SVMs," in Lecture Notes of Artificial

Intelligence 4114, D.-S. Huang, K. Li, and G. W.
Irwin, Eds. Springer, Berlin Heidelberg, 2006, pp. 701-710, doi:

10.1007/11816171_87. Computer Science & Information Technology (CS & IT) 377
https://ijisrt.com/assets/upload/files/IJISRT21JUN921.pdf

7. Imen Tayari Meftah,, Nhan Le Thanh & Chokri Ben Amar : Emotion Recognition Using KNN Classification for User Modeling and Sharing of Affect States.International Conference on Neural Information ProcessingCONIP 2012: **Neural Information Processing** pp 234–242

8. Schröder, M., Baggia, P., Burkhardt, F., Pelachaud, C., Peter, C., Zovato, E.: EmotionML – An Upcoming Standard for Representing Emotions and Related States. In: D'Mello, S., Graesser, A., Schuller, B., Martin, J.-C. (eds.)
ACII 2011, Part I. LNCS, vol. 6974, pp. 316–325. Springer, Heidelberg (2011)

9. Mohini Chaudhari and Sharvari Govilkar, Emotion Extraction using Rule based and SVM-KNN Algorithm.international Journal of Computer Applications (0975 – 8887) Volume 125 – No.11, September 2015

10. Priyanka Thakur and Dr. Rajiv Shrivastava. A review on Text Based Emotion
Recognition System, **International Journal of Advanced Trends in Computer Science and Engineering(2018)**

11. Text-based emotion detection: Advances, challenges, and opportunities:
https://onlinelibrary.wiley.com/doi/full/10.1002/eng2.12189

12. An unsupervised fuzzy clustering method for twitter sentiment analysis: https://ieeexplore.ieee.org/abstract/document/7779444

13. A novel intervention method for aspect-based emotion Using Exponential Linear Unit (ELU) activation function in a Deep Neural Network: https://ieeexplore.ieee.org/abstract/document/9432223

14. Fahad Mazaed Alotaibi ,Faculty of Computing and Information Technology in Rabigh (FCITR) King Abdul Aziz University (KAU) Jeddah Saudi Arabia https://www.academia.edu/2642268/CLASSIFICATION_OF_EMOTIONS_FROM_TEXT_USING_SVM_BASED

15. Text-Based Emotion Recognition Using Deep Learning Approach Santosh Kumar Bharti, S Varadhaganapathy, Rajeev Kumar Gupta, Prashant Kumar Shukla, Mohamed Bouye, Simon Karanja Hingaa , and Amena Mahmoud

16. Abdi A, Shamsuddin SM, Hasan S, Piran J (2019) Deep learning-based sentiment classification of evaluative text based on multi-feature fusion. Inf Process Manag 56(4):1245–1259

17. The HUMAINE Association: Humaine emotion annotation and representation language (earl): Proposal (Juin 2006), http://emotionresearch.net/projects/humaine/earl/proposal

18. Tayari, I., Le Thanh, N.: Sharing Emotional Information Using A Three Layer
Model. In: The Sixth International Conference on Internet and Web

Applications and Services, Netherlands Antilles, IARIA, Xpert Publishing Services (2011)

19. Schachter, S.: The interaction of cognitive and physiological determinants of emotional state. Advances in Experimental Social Psychology 1, 49–80 (1964)

20. Schachter, S.: The interaction of cognitive and physiological determinants of emotional state. Advances in Experimental Social Psychology 1, 49–80 (1964)

21. Ubeeka Jain and Amandeep Sandhu : A Review on the Emotion Detection from Text using Machine Learning Techniques, International Journal of Current Engineering and Technology, Accepted 16 July 2015, Available online 31 July 2015, Vol.5, No.4 (Aug 2015)