ISSN 2007-9737

Comparative Analysis of Emotion Detection Techniques Using Machine Learning

Sibani Patra[1], Rabi Narayan Satpathy[1], Chhabi Rani Panigrahi[2,*], Sarmistha Nanda[3]

[1] Sri Sri University, Odisha,
India

[2] Department of Computer Science,
Rama Devi Women's University,
India

[3] Department of Computer Applications,
ITER, S'O'A Deemed to be University,
India

{sibani.p2021-22ds, rabinarayan.s}@srisriuniversity.edu.in,
{panigrahichhabi, sarmisthananda}@gmail.com

Abstract. Emotion detection has emerged as a key focus area of research in human-computer interaction domain that allows systems to more effectively react to human emotional states. This paper presents a comparative analysis of various emotion detection techniques using machine learning (ML). We evaluate and compare the performance of different ML classifiers, including Bagging, K-Nearest Neighbor (KNN), and Random Forest (RF). The study involves the use of several publicly available datasets. The paper explores the benefits and constraints of each technique with respect to accuracy and computational efficiency, offering a comprehensive overview of current approaches in the field. The findings suggest that while RF algorithm outperforms KNN and Bagging techniques regarding accuracy. This work contributes valuable insights into selecting the most suitable emotion detection technique for different application scenarios. A comparative analysis of the classification algorithms is provided with the three considered datasets by computing the accuracy with recall, precision, and F-measure metrics. In the experimental results, RF algorithm archives a higher accuracy of 86%, 96%, and 85% in Dataset-1, Dataset-2, and Dataset-3 respectively which is better than the other considered algorithms.

Keywords. Machine learning, emotion detection, sentiment analysis, classifiers, random forest.

1 Introduction

Social media is one of the fastest-growing industries and the most widely used platform today. It has experienced rapid growth and has become a primary form of communication. Currently, there are over 5.17 billion active users, with a growth rate of 5.2% annually, equivalent to approximately 8.1 new users every second [11]. This significant rise of social media has the potential to affect people's lives in many ways. The data generated from these social media networks is vast and largely reflects our emotions, viewpoints, perspectives, and daily thoughts. Emotions play a crucial role in every human life.

Nowadays, users express their feelings in various formats, such as through posts, statuses, stories, and blogs, across platforms like Facebook, Instagram, Twitter, WhatsApp, Snap chat, etc. Detecting and understanding the emotions we experience daily is essential, as they influence our thoughts, actions, and decisions [1].

Emotions can be positive or negative. Positive emotions include happiness, joy, affection, care, and excitement, whereas negative emotions can be fear, anger, stress, depression, and anxiety [2]. Both positive and negative emotions should be balanced; otherwise it may lead an individual to face overwhelming challenges. Research findings indicate that there is a high connection between overuse of social media and an increased risk of depression, self-harming thoughts, loneliness, and anxiety.

The American Journal of Public Health [3] provides evidence that highlights this above said issue. The careful monitoring of online activities can offer valuable insights into individuals' behavior and personalities.

Therefore, creating systems that accurately recognize human emotions can improve the interaction between computers and humans. In addition to emotions detection, performing sentiment analysis can also help to identify a range of unique emotional responses, rather than simply categorizing them as negative, neutral, or positive. The ability to predict emotions has applications in various fields, including healthcare, stress monitoring, marketing, and customer satisfaction [10].

ML methods improve accuracy and reduce computational time as well as model development cost [9].

The paper organization is as follows. Section 2 describes the methodology used in this work and results are included in Section 3. In Section 4, a comparative analysis of the proposed work is presented; conclusion and future research directions are highlighted in Section 5.

2 Methodologies

We discuss in this section about the three datasets downloaded from Kaggle and the classification algorithms used in this work. We used Python programming version 3.10.12 for the implementation of ML classification algorithms in Google colab environment. The dataset was splitted in 80:20 ratio as training and testing sets.

2.1 Datasets Description

The datasets were obtained from Kaggle [7], an open source repository. The datasets Emotion dataset for NLP [5], Behavioral Tweets dataset [8], and Emotions in text [6] are named Dataset-1, Dataset-2 and Dataset-3, respectively in this work and the description of these datasets is provided in Table 1.

2.2 Classification Algorithms Used

The different features of these three classification algorithms considered, KNN, Bagging, and RF are given in Table 2.

3 Results and Findings

The results obtained from the experimentation are presented in this section along with the analysis of the considered classifiers. We considered three emotions datasets to evaluate the performance of each of the classifiers.

Three classifiers namely KNN, Bagging, and RF [4] were employed on these three datasets. We measured certain parameters such as the accuracy percentage; and the error metrics Mean Squared Error (MSE) and Root Mean Squared Error (RMSE), and the values of precision (P), recall (R), F-measure (F), and Support (S) with respect to accuracy metric. Using the weighted average and confusion matrix, we calculated all these metrics values for each of the ML algorithms.

The weighted average of accuracy of KNN, Bagging, and RF is given in Tables 3-5. The confusion matrix of only RF classifier is given as it gives the best results in all three datasets and is given as in Tables 6-8.

The comparative analysis of KNN, Bagging, and RF Algorithm is given in Table 9. Includes the accuracy, RMS, and RMSE for each of the classifiers in all three datasets.

Table 1. Datasets used in the study

Dataset Name	Emotions in text[8] (Dataset-1)	Behavioural Tweets Dataset[7] (Dataset-2)	Emotion dataset for NLP [6] (Dataset-3)
Instances	21405	7000	16000
Number of Emotions	06	02	06
Environment used	Google colab (python 3.10.12)	Google colab (python 3.10.12)	Google colab (python 3.10.12)
File Type	.CSV	.CSV	.CSV

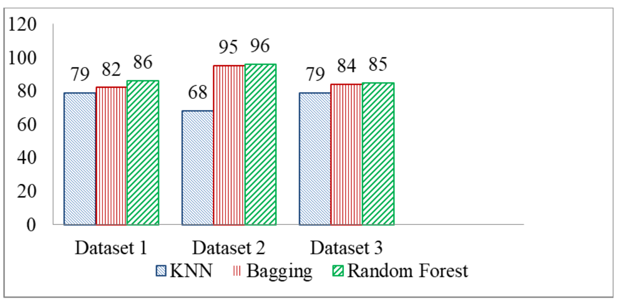

Fig. 1. Comparative analysis in terms of accuracy of of KNN, Bagging, and RF algorithms

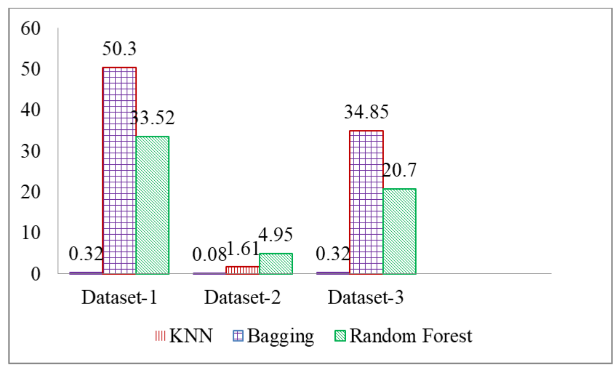

Fig. 2. Comparative analysis in terms of model building time of KNN, Bagging, and RF algorithms

4 Comparative Analysis

The accuracy comparison results are shown in Figure 1 for the KNN, Bagging, and RF algorithms which were applied to all the three considered datasets. From experimental results, it is concluded that RF algorithm gives higher accuracy in all the three datasets with regard to other algorithms in terms of accuracy. The accuracy obtained for RF algorithm in three considered datasets were 86%, 96%, and 95% respectively and is found to be higher in RF as compared to KNN and Bagging. The model building time of KNN algorithm is less in all the three datasets as compared to Bagging and RF algorithm. The model building time was found to be 0.32s, 08s, and 0.32s for Dataset-1, Dataset-2, and Dataset-3 respectively and is displayed as in Figure 2. The MSE as well as RMSE were also calculated for all algorithms in the three datasets. The MSE and RMSE values were found to be higher for KNN as compared to Bagging and RF.

5 Conclusion and Future Work

In this work a comparative analysis was performed to evaluate the performance of the three classification algorithms, namely KNN, Bagging, and RF with regard to metrics such as P, R, F, and S. From the experimental results, we conclude that the considered classification algorithms are suitable to classify human emotions. Recognizing human emotions is useful and important for various real life applications. This study can also help researchers in making decisions on applying classification algorithms for detection of human emotions.

Table 2. Comparison of Machine Learning Algorithms

Algorithm/Feature	**KNN**	**Bagging**	**Random Forest (RF)**
Type	Instance-based (lazy learning)	Ensemble method	Ensemble method
Principle	It finds the 'k' closest training examples in the feature space and predicts the class based on the average of those neighbours or majority vote.	The predictions of these models are combined by averaging (for regression) or voting (for classification) to improve accuracy and reduce variance.	Multiple decision trees are built during training and output the class that is the average of the individual trees' predictions or the majority vote. It introduces randomness by selecting random subsets of features for each tree, reducing overfitting and improving generalization.
Model building	There is no explicit training phase in KNN. During prediction, the algorithm searches the training data to find the conclusión final.	Bagging involves training multiple instances of the same base model (e.g., decision trees) on bootstrapped subsets of the data and combining their outputs.	It creates many decision trees, each trained on a bootstrapped subset of the data. It then aggregates the predictions from all trees to make the final decision.
Complexity	High computational cost at prediction time, as it needs to compute distances to all training points. Training is fast, but testing is slow with large datasets.	Requires significant computational resources during training due to the multiple models being trained, but prediction is faster compared to deep learning.	Training can be computationally intensive due to the need to build multiple decision trees. However, prediction time is relatively fast as it involves averaging or voting among trees.
Advantage	Easier to understand and implement. Gives good results with small datasets where relationships between instances are clear.	Reduces variance and prevents overfitting, especially for high-variance models like decision trees.	Reduces overfitting compared to single decision trees, handles missing data well, and is robust to outliers.
Disadvantage	Computationally expensive for large datasets, and performance can degrade with high-dimensional data.	Can be computationally expensive for large datasets. Not effective at reducing bias if the base model is biased.	May be computationally expensive in terms of memory and time, especially with many trees and large datasets.

The findings of this comparative analysis of emotion detection techniques offer several directions for future research. One important direction is the exploration of hybrid models that

ISSN 2007-9737

Table 3. Weighted Average of Accuracy of the KNN Algorithm

Dataset	Emotions	P	R	F	S
Dataset-1	anger	.75	.80	.78	617
	fear	.79	.73	.76	531
	happy	.79	.87	.83	1381
	love	.71	.57	.63	318
	sadness	.85	.83	.84	1277
	surprise	.71	.48	.57	168
Dataset-2	Anxious	.61	.99	.76	695
	stressed	.99	.37	.54	705
Dataset-3	anger	.74	.81	.77	427
	fear	.75	.71	.73	397
	joy	.78	.88	.83	1021
	Love	.78	.50	.61	296
	sadness	.83	.84	.84	946
	surprise	.72	.43	.54	113

Table 4. Weighted Average of Accuracy of the Bagging Algorithm

Dataset	Emotions	P	R	F	S
Dataset-1	anger	.83	.84	.84	617
	fear	.81	.85	.83	531
	happy	.81	.86	.83	1381
	love	.71	.62	.66	318
	sadness	.88	.84	.86	1277
	surprise	.77	.65	.71	168
Dataset-2	Anxious	.93	.98	.95	695
	stressed	.98	.92	.95	705
Dataset-3	anger	.81	.89	.85	427
	fear	.83	.85	.84	397
	joy	.83	.88	.85	1021
	Love	.79	.63	.70	296
	sadness	.90	.86	.88	946
	surprise	.71	.66	.68	113

Table 5. Weighted Average of Accuracy of the RF Algorithm

Dataset	Emotions	P	R	F	S
Dataset-1	anger	.88	.86	.87	617
	fear	.82	.85	.83	531
	happy	.84	.92	.87	1381
	love	.80	.67	.73	318
	sadness	.91	.88	.89	1277
	surprise	.82	.63	.71	168
Dataset-2	Anxious	.93	.99	.96	695
	stressed	.99	.93	.96	705
Dataset-3	anger	.84	.82	.83	427
	fear	.85	.82	.84	397
	joy	.81	.93	.87	1021
	Love	.82	.63	.71	296
	sadness	.92	.88	.90	946
	surprise	.78	.67	.72	113

Table 6. Confusion Matrix of RF for Dataset-1

Emotions	anger	fear	happy	love	sadness	surprise
anger	532	20	30	3	30	2
fear	18	449	25	2	27	10
happy	24	16	1265	33	37	6
love	3	4	90	212	9	0
sadness	29	30	81	13	1119	5
surprise	2	31	20	1	8	106

Table 7. Confusion Matrix of RF for Dataset-2

Emotions	anxious	stressed
anxious	686	9
stressed	52	653

combine the strengths of both traditional ML algorithms and deep learning approaches.

The accuracy and efficiency of emotion detection, particularly in multimodal contexts may be obtained by integrating these techniques.

Additionally, we will plan to investigate the impact of advanced pre-processing techniques, such as emotion-specific feature extraction, to improve model performance. The integration of real-time emotion detection in dynamic environments, such

Table 8. Confusion Matrix of RF for Dataset-3

Emotions	anger	fear	joy	love	sadness	surprise
anger	350	13	32	3	29	0
fear	16	326	26	2	15	12
joy	11	8	949	28	20	5
love	5	6	93	187	5	0
sadness	32	14	56	7	833	4
surprise	1	15	13	1	7	76

Table 9. Comparative analysis of KNN, Bagging, and RF Algorithms

Dataset	Algorithm	Accuracy	MSE	RMSE
Dataset-1	KNN	79	1.227	1.107
	Bagging	82	0.995	0.997
	Random Forest	**86**	0.804	0.897
Dataset-2	KNN	68	0.318	0.564
	Bagging	95	0.046	0.215
	Random Forest	**96**	0.043	0.208
Dataset-3	KNN	79	1.188	1.09
	Bagging	84	0.94	0.969
	Random Forest	**85**	0.827	0.909

as Augmented Reality (AR) or Virtual Reality (VR) presents another area for research.

Competing Interests

The authors have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

References

1. **An, S., Ji, L.-J., Marks, M., Zhang, Z. (2017).** Two sides of emotion: Exploring positivity and negativity in six basic emotions across cultures. frontiers in psychology. Vol. 8, pp. 253368.
2. **Cherry, K. (2022).** The important role of emotions.
3. **Dean, B. (2021).** Social network usage growth statistics: How many people use social media in 2021.
4. **Díaz-Ramírez, A., Diaz-Escobar, J., Quintero-Rosas, V., Moncada-Sanchez, R. (2024).** Classification of fall events in the elderly using a thermal sensor and machine learning techniques. DOI: 10.13053/CyS-28-4-4809.
5. **Govi, P. (2025).** Emotions dataset for NLP.
6. **Juyal, I. (2025).** Emotions in text dataset.
7. **Kaggle Contributors (2025).** Kaggle datasets.
8. **Kandroo, A. (2025).** Behavioural tweets dataset.
9. **Nanda, S., Panigrahi, C., Pati, B. (2023).** Flood prediction with optimized attributes and clustering. Computación y Sistemas, Vol. 27, No. 4, pp. 1157–1167. DOI: 10.13053/CyS-27-4-4643.
10. **Nasim, S. (2025).** Speech emotion recognition using deep learning.
11. **Robinson, L., Melinda Smith, M. (2025).** Social media and mental health: Social media addiction.

Article received on 18/04/25; accepted on 10/06/2025.
**Corresponding author is Chhabi Rani Panigrahi.*