Cloud Computing and Data Science
https://ojs.wiserpub.com/index.php/CCDS/

Research Article

Developing an Emotion Recognition Tool for Tweets

Harshit D. Bhavnani

Department of Data Science, Mukesh Patel School of Technology Management & Engineering, Narsee Monjee Institute of Management Studies University, Mumbai, Maharashtra, India
E-mail: harshit.bhavnani@gmail.com

Received: 19 August 2022; **Revised:** 18 November 2022; **Accepted:** 29 November 2022

Abstract: This research in the domain of sentiment analysis aims to maintain high effectiveness and applicability by not just understanding how the user feels, but by accurately quantifying their emotions. With the desired end product being a web application, the visualization that is received as an output aids the user in comparing the intensity of their emotions of anger, fear, joy and sadness demonstrated in the input tweets. This web application was built with the use of a dataset from the SEMEVAL-2018 competition. Training and testing of the dataset using 6 machine learning algorithms and their evaluation using performance metrics including R2, MAE, MSE and RMSE led us to arrive at the result that 'Support Vector Regressor' was the best performing algorithm for anger, sadness and fear while the Gradient Boosting algorithm performed best for joy. To that end, the web application uses the following algorithm for evaluating the respective emotions. In addition to developing a prediction model, the research also involved extensive data visualization and analysis that conveyed the most used words and hashtags when the user experiences each of the aforementioned emotions. To highlight the research's accomplishment-the author has been able to create a fairly accurate and relatively quick sentiment analysis model for public opinion.

Keywords**:** emotion recognition, natural language processing, machine learning, data analysis

1. Introduction

In times when mental health has been getting its due importance by society and when people are understanding the need of assessing one's mental health regularly and taking corrective action for it, we have realized that not every individual may possess enough emotional intelligence [1]. Thus, a tool that accurately evaluates and effectively represents our sentiments would add value to the world, which is the motivation behind conducting this research that led to the development of our 'Emotion Recognition Tool'.

An individual's feelings are most of the time influenced by the ecosystem they live in. Social media, especially Twitter, is widely used as a platform to express one's thoughts and opinions to the world and on most occasions, users are often candid [2]. Thus, in order to understand how an individual feels, the opinions they share on Twitter could be analyzed to evaluate their sentiments.

This tool accepts tweets made by users, analyzes the intensity of anger, joy, sadness and fear in those tweets and displays a bar plot that mentions the percentage share of the aforementioned sentiments in the tweet. Applications of this

DOI: https://doi.org/10.37256/ccds.4120231827

tool in the real world are endless, ranging from stock trading to security practices. [A copy of the dataset, source code and an informative presentation can be accessed using the following link: https://github.com/harshitbhavnani/Emotion-Recognition-for-Tweets].

Our goal with this research is to present a fairly accurate and relatively quick sentiment analysis model for public opinion to a community of budding researchers, which could be used by them for further exploration into the field of sentiment analysis.

2. Literature review

While working on the same data, Mohammad and Bravo-Marquez [3] used a technique called Best-Worst Scaling (BWS) which improved their annotation intensity and gave them better results [4]. They also found out that hashtags usually convey intense emotions. In addition to all this, they evaluated feature importance through a benchmark regression system and the similarity between emotions.

Chiorrini et al. [5] make use of Bidirectional Encoder Representations from Transformers (BERT) models for both sentiment analysis and emotion recognition of Twitter data.

In [6] the authors combined lexical, syntactic and pre-trained word embedding features and used AdaBoost with XGBoost as the base regressor. For hyper-parameter tuning, 10-fold cross-validation with optimization Pearson's correlation score was used.

3. Data

The dataset used has 4 columns - ID, Tweet, Affect Dimension and Intensity Score. The 'ID' column contains the personal unique IDs of users and thus, wasn't useful in order to conduct research or develop the web application [7]. The 'Tweet' column consisted of the tweets made by users. The 'Affect Dimension' column consisted of the respective emotion out of the four basic emotions: anger, fear, joy, and sadness that is demonstrated through the corresponding tweet and the 'Intensity Score' column consisted of the quantified value (between 0 to 1) of the sentiment felt by the user. The language of the data derived from our source is English. A sample structure for the dataset is illustrated in Table 1.

Table 1. Sample structure of the dataset

ID	Tweet	Affect dimension	Intensity score
2017-En-30692	Positive research show salesperson score top g ...	joy	0.274
2017-En-11102	passed away early morning fast furious styled ...	anger	0.354
2017-En-41401	If Troyler die Im gonna die	sadness	0.798
2017-En-21664	terrorism booming industry Pak govt oblivious ...	fear	0.625

4. Methodology

The pipeline of this research could be divided broadly into 4 parts - Increasing quality and preparing the data for analysis and prediction, understanding the data and deriving insights from it, finding the best prediction algorithms for the sentiment intensity scores and interpreting the results for their implementation on the end product.

4.1 *Cleaning*

Each key was split into tokens and the tokens were deleted if they consisted of:
• Tagged Accounts
• Links
• Removing Empty/Duplicate Tweets
• Punctuation marks and Special Characters
• Stopwords (Standard English Stopwords from the NLTK Library)
• NA Values

4.2 *Preparation*

To prepare the data for processing, feature extraction is done on the dataset which extracts key phrases that could be used for modeling and development. Feature extraction is followed by word2vec embedding using the google news collection.

4.2.1 *Feature extraction*

The Textblob Conll Extractor, a noun phrase extractor that uses chunk parsing trained with the ConLL-2000 training corpus is used for extracting keyphrases from the tweets.

4.2.2 *Embedding*

Word2vecs [8] embedding trained on Google News collection, which has almost become a standard embedding for research focusing on natural language processing, has been used for this project as well. This embedding is essentially a 300-dimensional vector built using the Google News dataset of 100 billion tokens and 3 million distinct words and phrases.

4.3 *Exploratory data analysis*

To gain additional insights from the data, we create word clouds of each emotion to receive the most used words for each sentiment. Additionally, bar plots are created to represent the most frequently used hashtags for each emotion.

Figure 1. Word cloud depicting most used words in anger

Figure 2. Word cloud depicting most used words in joy

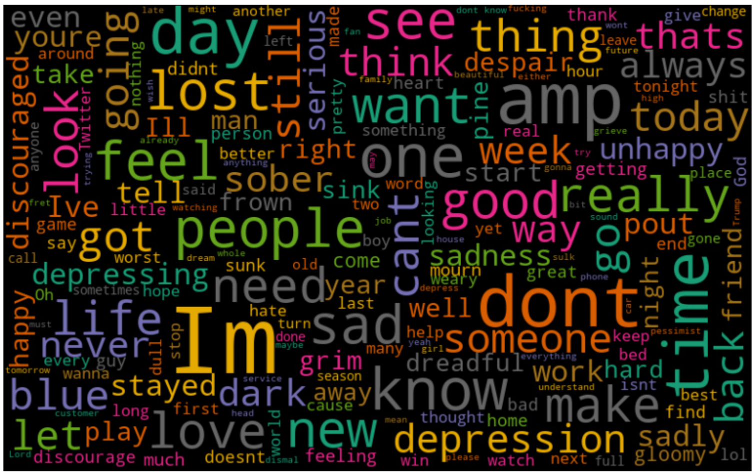

Figure 3. Word cloud depicting most used words in sadness

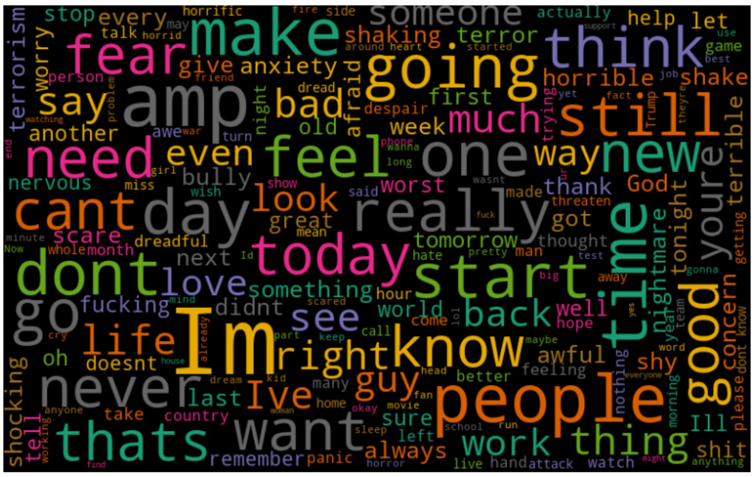

Figure 4. Word cloud depicting most used words in fear

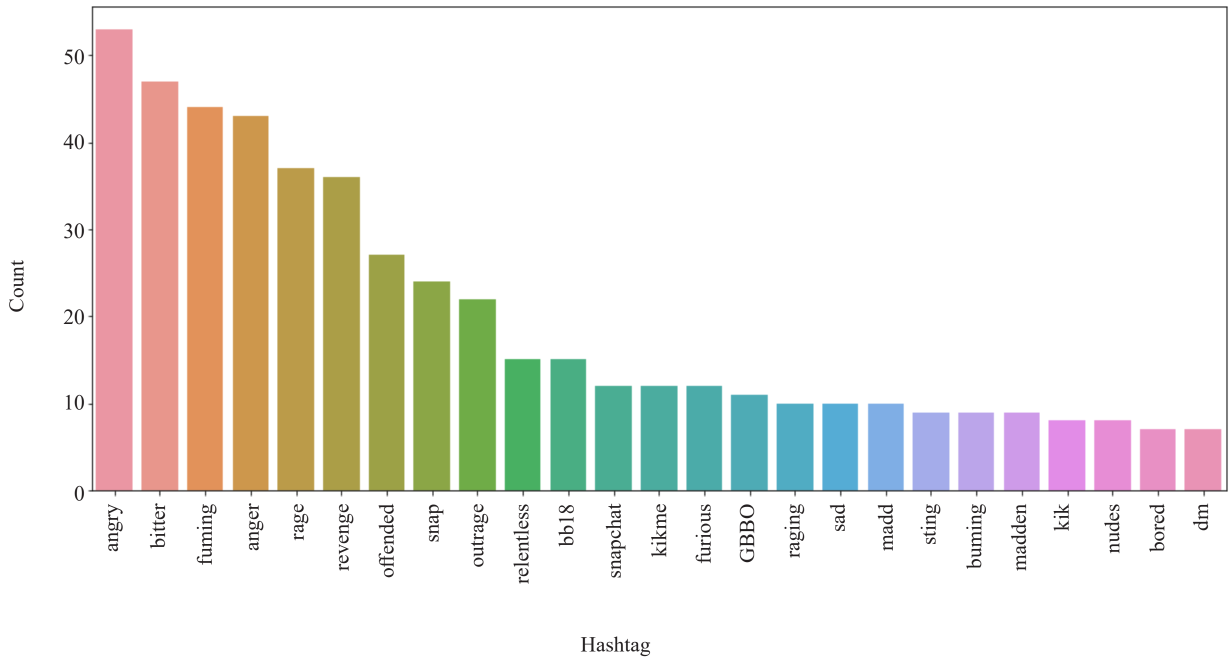

Figure 5. Bar Plot representing most used hashtags in anger

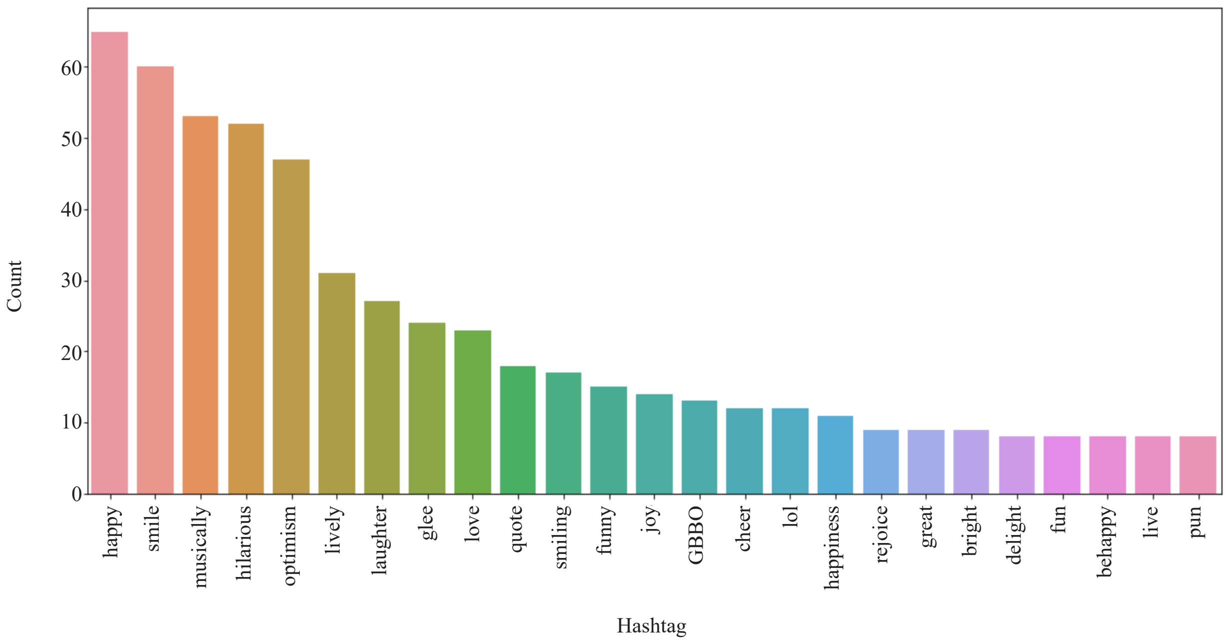

Figure 6. Bar Plot representing most used hashtags in joy

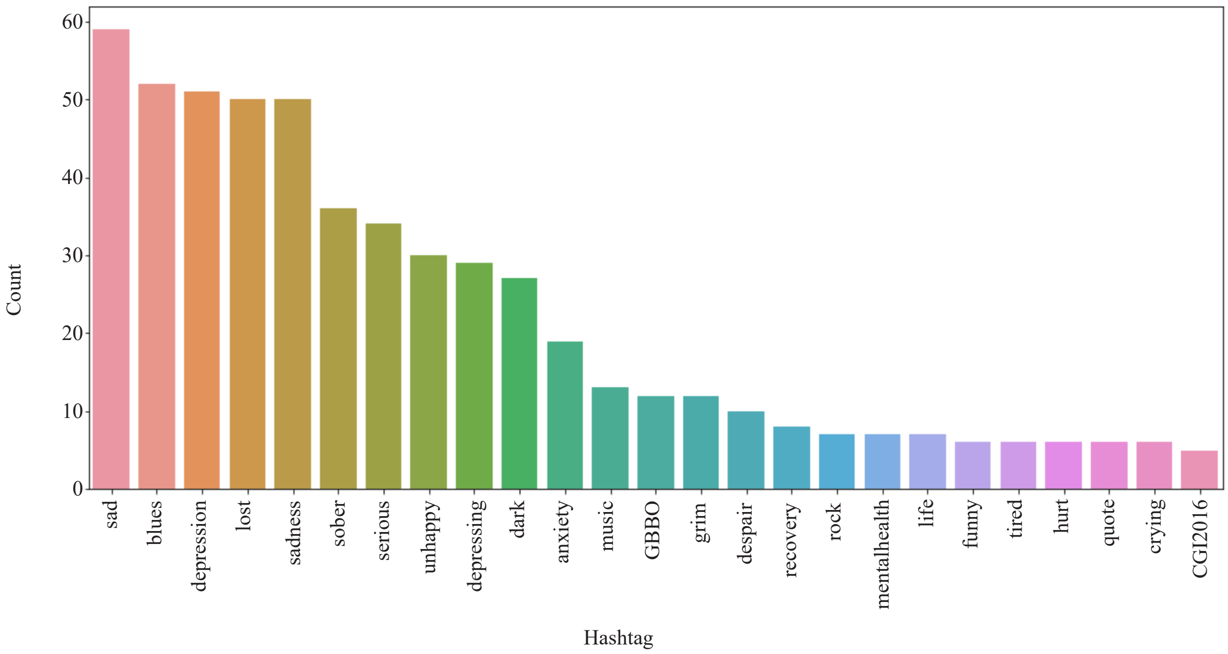

Figure 7. Bar Plot representing most used hashtags in sadness

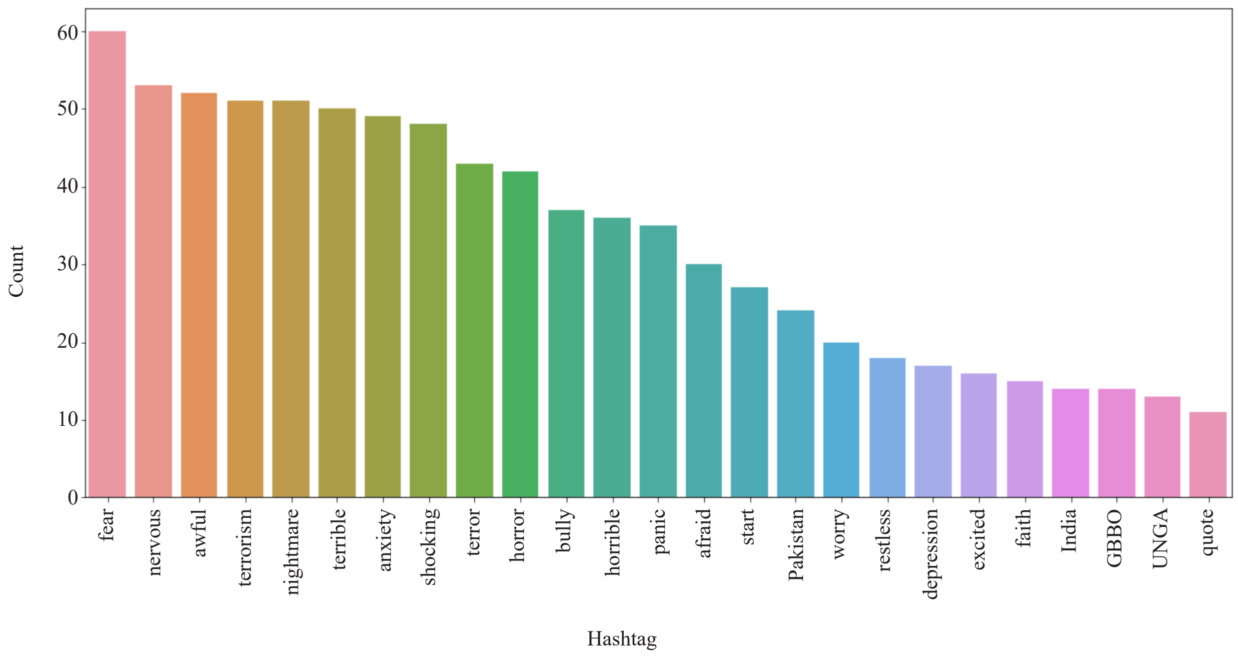

Figure 8. Bar Plot representing most used hashtags in fear

4.4 *Experimental study*

Six algorithms are evaluated using 4 performance metrics to find the best-suited algorithm for each of the four sentiments' intensity scores.

4.4.1 *Algorithms*

This subsection explains the motivation behind using certain statistical techniques. Most of them were adopted after conducting the literature review and observing that they perform well in similar cases.

4.4.1.1 *Linear regression*

Ginosar Shiry and Steinitz Avital [9] in their research have stated that although linear regression is simple, it performs fairly well.

4.4.1.2 *Random forest regressor*

In the research carried out by Arnav Munshi et al. [10], they used the random forest regressor for conducting sentiment analysis on tweets which gave a significant accuracy to the results. Saifullah Shoffan et al. [11] have also stated that random forest was the best algorithm for their analysis. Khanvilkar Gayatri and Vora Deepali [12] also concluded that random forest works best for their data.

4.4.1.3 *Decision tree regressor*

Aufar Mohammad et al. [13], in their research, compared the performance of the decision tree algorithm with the random forest algorithm on data extracted from YouTube where the Decision Tree algorithm had a slightly higher accuracy than the Random Forest algorithm.

4.4.1.4 *K-nearest neighbors regressor*

Huq Rezwanul Mohammad et al. [14] use K-Nearest Neighbor and support vector machine algorithm for conducting sentiment analysis on tweets.

4.4.1.5 *Support vector regressor*

Korovkinas Konstantinas and Garsva Gintautas [15], in their research in the field of sentiment analysis concluded Logistic Regression and SVM, and a combination of these methods fit the best for their further work in the field of sentiment analysis.

4.4.1.6 *Gradient boosting*

While developing a predictive model for sentiment analysis of discussion or review content on social media, Kumar Pradeep and Wahid Abdul [16] stated that their experimental results revealed that XGBM outperforms in terms of both training and testing accuracy.

4.4.2 *Performance metrics*

The four performance metrics used are R2, MAE, MSE and RMSE. The scores achieved in terms of the coefficient of determination are given more importance as it conveys the ability of how well the model fits the dependent variables.

4.4.2.1 *Coefficient of determination*

The Coefficient of Determination, also known as R squared is the proportion of the variation in the dependent variable that is predictable from the independent variable(s). It is the square of the Correlation Coefficient(R). Although R2 does not account for overfitting, as there is less number of independent variables, it can be used.

$$R^2 = 1 - \frac{\sum (y_i - \hat{y})^2}{(y_i - \overline{y})^2}$$

4.4.2.2 *Mean absolute error*

The Mean absolute error represents the average of the absolute difference between the actual and predicted values in the dataset. It measures the average of the residuals in the dataset.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}\right|$$

4.4.2.3 *Mean squared error*

Mean Squared Error represents the average of the squared difference between the original and predicted values in the data set. It measures the variance of the residuals.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2$$

4.4.2.4 *Root mean squared error*

Root Mean Squared Error is the square root of Mean Squared error. It measures the standard deviation of residuals.

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2}$$

4.5 *Results & analysis*

Performance of all the algorithms while predicting sentiment intensity scores is listed in the tables below (Table 2-5):

Table 2. Performance of machine learning algorithms on tweets of anger

	Anger					
	Linear regression	Random forest	Decision tree	KNN	SVR	GradientBoost
R2	-7.3828e + 21	-0.0788643	-0.182579	-0.140585	-0.0153836	-0.0223897
MAE	3.48474e + 09	0.170682	0.177707	0.17367	0.165721	0.16674
MSE	2.92349e + 20	0.0427216	0.0468286	0.0451657	0.0402079	0.0404853
RMSE	1.70982e + 10	0.206692	0.216399	0.212522	0.200519	0.20121

Table 3. Performance of machine learning algorithms on tweets of joy

	Joy					
	Linear regression	Random forest	Decision tree	KNN	SVR	GradientBoost
R2	-1.96711e + 22	-0.0910428	-0.159834	-0.384159	-0.0149771	-0.0109779
MAE	3.18182e + 09	0.162775	0.167744	0.182052	0.161567	0.159749
MSE	7.22398e + 20	0.0400672	0.0425935	0.0508316	0.0372738	0.0371269
RMSE	2.68775e + 10	0.200168	0.206382	0.225459	0.193064	0.192684

Table 4. Performance of machine learning algorithms on tweets of sadness

	Sadness					
	Linear regression	Random forest	Decision tree	KNN	SVR	GradientBoost
R2	-6.71138e + 21	-0.106157	-0.207057	-0.236501	-0.00589185	-0.0225636
MAE	3.14869e + 09	0.153242	0.159843	0.161546	0.146084	0.147382
MSE	2.18221e + 20	0.0359667	0.0392475	0.0402049	0.0327066	0.0332487
RMSE	1.47723e + 10	0.189649	0.19811	0.200511	0.18085	0.182342

Table 5. Performance of machine learning algorithms on tweets of fear

	Fear					
	Linear regression	Random forest	Decision tree	KNN	SVR	GradientBoost
R2	-5.0432e + 21	-0.161456	-0.227915	-0.209392	-0.0226743	-0.0947355
MAE	2.50807e + 09	0.151374	0.155281	0.156681	0.144375	0.147129
MSE	1.58992e + 20	0.036616	0.0387112	0.0381272	0.0322408	0.0345126
RMSE	1.26092e + 10	0.191353	0.196752	0.195262	0.179557	0.185776

Although Gradient Boosting Regressor performs well, Support Vector Regressor performs slightly better for predicting anger, sadness and fear while Gradient Boosting Regressor works slightly better than the Support Vector

Regressor for predicting joy. A notable observation is that Linear Regression gives extremely poor results and should be avoided while predicting sentiments. Additionally, the random forest had good reviews, but it did not perform well for our data.

5. Web application

The end product of this research is a Flask-based Web Application which is hosted by ngrok. The users are required to enter a tweet. The application extracts, transforms and loads the data entered by the user into the pre-trained model which makes predictions using the best performing algorithms for each sentiment and the results obtained are represented in the form of a bar plot mentioning the percentage share of each emotion in their tweet.

5.1 *Sample input*

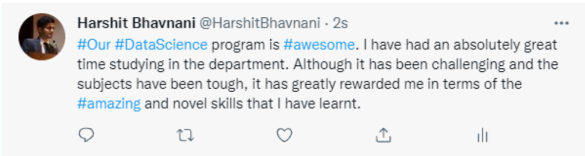

Figure 9. Sample input (Tweet)

5.2 *Sample output*

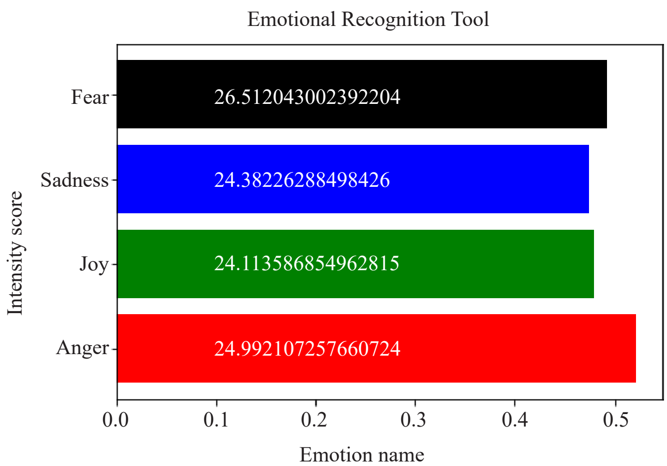

Figure 10. Sample output (Horizontal Bar Plot)

6. Conclusion & future scope

The results are easily understandable and fairly accurate which makes this project highly scalable. The application of this project in the psychological/psychiatric space would aid the industry greatly. It would also help researchers build upon the findings and develop an algorithm that gives a detailed and extensively elaborate output and help the users in a better manner.

A restriction that we faced while conducting this research was that it was primarily done with the purpose to build a web-application. Thus, the use of time-consuming deep learning algorithms or complex embedding techniques such as Stanford's GloVe [17]. This could be overcome by researchers who would probe further into emotion recognition.

Conflicts of interest

The author declares no competing financial interest.

References

[1] World Health Organization. *Investing in Mental Health*. n.d. Available from: https://www.who.int/mental_health/media/investing_mnh.pdf [Accessed 23th March 2022].

[2] Utz S. Social media as sources of emotions. In: Sassenberg, K., Vliek, M. (eds.) *Social Psychology in Action*. Springer, Cham; 2019. pp. 205-219. Available from: doi: 10.1007/978-3-030-13788-5_14.

[3] Mohammad SM, Bravo-Marquez F. Emotion intensities in Tweets. In *Proceedings of the 6th Joint Conference on Lexical and Computational Semantics (*SEM 2017)*. Vancouver, Canada: Association for Computational Linguistics; 2017. pp. 65-77. Available from: doi: 10.18653/v1/S17-1007.

[4] Louviere J, Flynn T, Marley A. *Best-Worst Scaling: Theory, Methods and Applications*. Cambridge: Cambridge University Press; 2015. Available from: doi: 10.1017/CBO9781107337855.

[5] Chiorrini A, Diamantini C, Mircoli A, Potena D. *Emotion and sentiment analysis of tweets using BERT*. 2021. Available from: http://ceur-ws.org/Vol-2841/DARLI-AP_17.pdf [Accessed 23th March 2022].

[6] Duppada V, Sushant H. Seernet at emoInt-2017: Tweet emotion intensity estimator. *Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*. Association for Computational Linguistics; 2017, pp. 205-211. Available from: doi: 10.18653/v1/W17-5228.

[7] Saif MM, Felipe B-M, Mohammad S, Svetlana K. In *Proceedings of the International Workshop on Semantic Evaluation (SemEval-2018)*. New Orleans, LA, USA; 2018. Available from: http://saifmohammad.com/WebDocs/semeval2018-task1.pdf [Accessed 23th March 2022].

[8] Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J. Distributed representations of words and phrases and their compositionality. In Burges CJC, Bottou L, Welling M, Ghahramani Z, Weinberger KQ (eds.) *Advances in Neural Information Processing Systems*. Opgehaal van; 2013. Available from: https://proceedings.neurips.cc/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf [Accessed 23th March 2022].

[9] Ginosar S, Steinitz A. *Sentiment analysis using linear regression*. 2012. Available from: http://bid.berkeley.edu/cs294-1-spring12/images/8/89/ShiryGinosar_AvitalSteinitz2.pdf [Accessed 23th March 2022].

[10] Munshi A, Sapra S, Arvindhan M. A novel random forest implementation of sentiment analysis. 2020. Available from: https://www.irjet.net/archives/V7/i6/IRJET-V7I6532.pdf [Accessed 23th March 2022].

[11] Shoffan S, Yuli F, Agus SA. *Comparison of machine learning for sentiment analysis in detecting anxiety based on social media data*. 2021. Available from: doi: 10.48550/ARXIV.2101.06353.

[12] Khanvilkar G, Vora D. Product recommendation using sentiment analysis of reviews: A random forest approach. *International Journal of Engineering and Advanced Technology (IJEAT)*. 2022; 8(2s2): 146-152. Available from: https://www.ijeat.org/wp-content/uploads/papers/v8i2s2/B10320182S219.pdf [Accessed 24th March 2022].

[13] Aufar M, Andreswari R, Pramesti D. Sentiment analysis on Youtube social media using decision tree and random forest algorithm: A case study. *2020 International Conference on Data Science and Its Applications (ICoDSA)*. 2020. pp. 1-7. Available from: doi: 10.1109/ICoDSA50139.2020.9213078.

[14] Huq MR, Ali A, Rahman A. *Sentiment analysis on Twitter data using KNN and SVM*. 2017. Available from: https://thesai.org/Downloads/Volume8No6/Paper_3-Sentiment_Analysis_on_Twitter_Data_using_KNN_and_SVM.pdf [Accessed 23th March 2022].

[15] Korovkinas K, Garsva G. *Selection of intelligent algorithms for sentiment classification method creation*. n.d. Available from: http://ceur-ws.org/Vol-2145/p26.pdf [Accessed 23th March 2022].

[16] Kumar P, Wahid A. *Social media analysis for sentiment classification using gradient boosting machines*. 2021. Available from: https://easychair.org/publications/preprint_open/2bxD [Accessed 23th March 2022].

[17] Pennington J, Socher R, Manning CD. *GloVe: Global Vectors for Word Representation*. n.d. Available from: https://nlp.stanford.edu/pubs/glove.pdf [Accessed 23th March 2022].