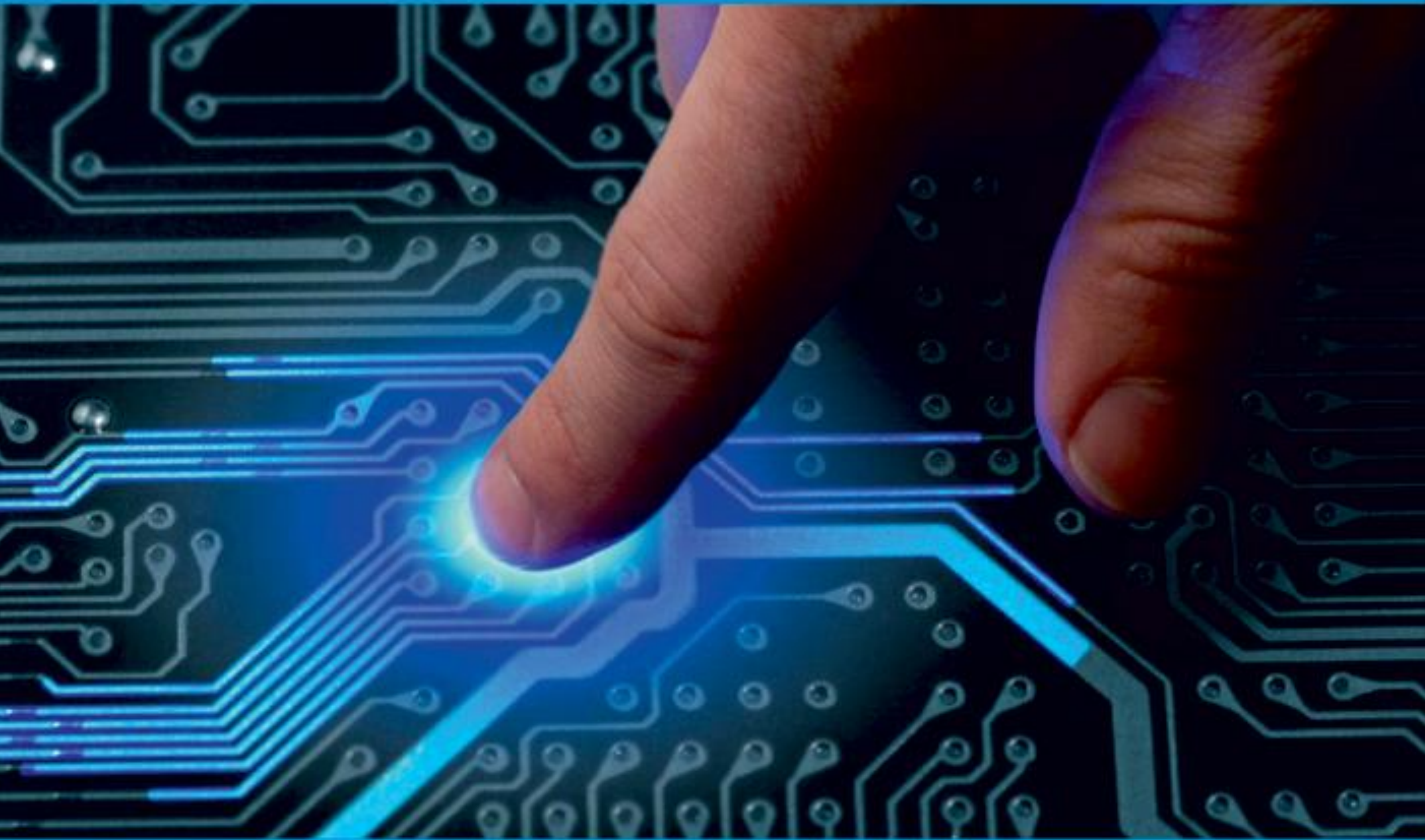
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 4, April 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 4, April 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1004153|

Emotion Recognition from Twitter

Bhukan Pratiksha Popat, Date Siddhesh Bhau, Hase Shubhangi Bharat, Wakchaure Kajal Ashok, Dr. Borkar B. S.

Department of Information Technology, AVCOE, Ahmednagar, Maharashtra, India

Department of Information Technology, AVCOE, Ahmednagar, Maharashtra, India

Department of Information Technology, AVCOE, Ahmednagar, Maharashtra, India

Department of Information Technology, AVCOE, Ahmednagar, Maharashtra, India

Department of Information Technology, AVCOE, Ahmednagar, Maharashtra, India

ABSTRACT: Twitter is an online social networking service that has more than 300 million users, generating a huge amount of information each day. Twitter's most important feature is its ability for users to tweet about events, situations, feelings, opinions, or even something new in real-time. There is no system of accuracy or reliability in place. Anyone can say just anything and express their feeling through tweets. These tweets are interrelated to emotions of individual. The emotions expressed can be as happy as some achievements and as dangerous as last note of suicide or shamming someone. To recognize the same the system is proposed for emotion recognition of tweets. For this use of hashtags is made to create three large emotion labelled data sets corresponding to different classifications of emotions. Comparison of several words and character-based recurrent and convolutional neural networks with the performance on bagof-words and latent semantic indexing models is performed. Emotions are recognized through application and analysis of the transferability of the final hidden state representations between different classifications of emotions is shown with recurrent neural networks, especially about character-based ones are better than bag-of-words and latent semantic indexing models.

KEYWORDS: Agriculture , Android Application , Machine Learning , IoT , ChatBot

I. INTRODUCTION

The development of social network platforms has given people a new way to generate and consume a great deal of information on the web. In the past, people used to get information from portal websites. A large number of websites provide a long list of topics varying from politics to entertainment. These traditional online information sources are useful but less efficient because they often contain redundant information. However, since the arrival of online social network platforms, people tend to get information from these platforms because of their fast and efficient features. These platforms are available for users to choose the information source they are interested in. And also a large number of social network platforms such as Twitter, Google+, and Facebook provide information for users. Twitter is the most popular microblogging platform in the world. It is also the fastest growing social network platform and has a dominant position in the area of mircroblogging. More than 500 million registered users post 340 million twitter messages every day, sharing their opinions and daily activities. Compared with regular microblogging platforms, Twitter messages are much shorter. You are only allowed to post 140 characters or less in one Twitter message. This feature makes Twitter easier for people to get the main point from the massive amount of information available online. Depending on the need of the users, Twitter users can follow whichever people and information source they prefer. With all of the advantages mentioned above, Twitter thus has become a powerful platform with many kinds of information from worldwide breaking news to purchasing products at home.

The experiment conducted as part of this experiment uses datasets of tweets. Due to inherent domain specific nature of the sentiment and emotion analysis, applying this study to other datasets in different social media services can yield different results and new insights. The main focus of this study is evaluating the accuracy of the proposed approach of emotion analysis.

Paper is organized as follows. Section II describes about the related work done earlier for the system to be developed. Section III presents method used and algorithms used for the detection. Section IV presents experimental results showing results of images tested. Finally, Section V presents conclusion.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 4, April 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1004153|

II. RELATED WORK

• Matla Suhasini, Srinivasu Badugu, "Two Step Approach for Emotion Detection on Twitter Data", International Journal of Computer Applications, June 2018

"Emotional states of individuals, also known as moods, are central to the expression of thoughts, ideas and opinions, and in turn impact attitudes and behavior". In this paper we have proposed a method which detects the emotion or mood of the tweet and classify the twitter message under appropriate emotional category. Our approach is a two-step approach, it is so called as it uses two approaches for the classification process, one is Rule Based approach and the other is Machine Learning approach. The first approach is the Rule Based Approach (RBA), our minor contributions in this approach are pre-processing, tagging, feature selection and Knowledge base creation. Feature selection is based on tags. Our second approach is Machine Learning Approach (MLA), in this the classifier is based on supervised machine learning algorithm called Na¨ıve Bayes which requires labeled data. Na¨ıve Bayes is used to detect and classify the emotion of a tweet. The output of RBA is given to MLA as input because MLA requires labeled data which we have already created through RBA. We have compared the accuracies of both the approaches, observed that, with the rule based approach we are able to classify the tweets with accuracy around 85% and with the machine learning approach the accuracy is around 88%. Machine learning approach performance is better than rule based approach, the performance has been improved as we have removed the error data while training the model. The approaches are involved with the concepts of Natural Language Processing, Artificial Intelligence, and Machine Learning for the development of the system. Our major contributions in this paper are detection of emotion for non hashtagged data and the labeled data creation for machine learning approach without manual creation. [1]

•

• Srinivasu Badugu, Matla Suhasini, "Emotion Detection on Twitter Data using Knowledge Base Approach", International Journal of Computer Applications, March 2017

Emotional states of individuals, also known as moods, are central to the expression of thoughts, ideas and opinions, and in turn impact attitudes and behavior. Social media tools like twitter is increasingly used by individuals to broadcast their day-to-day happenings or to report on an external event of interest, understanding the rich 'landscape' of moods will help us better to interpret millions of individuals. This paper describes a Rule Based approach, which detects the emotion or mood of the tweet and classifies the twitter message under appropriate emotional category. The accuracy with the system is 85%. With the proposed system it is possible to understand the deeper levels of emotions i.e., finer grained instead of sentiment i.e., coarse grained. Sentiment says whether the tweet is positive or negative but the proposed system gives the deeper information of tweet which has adverse uses in the field of Psychology, Intelligence Bureau, Social and Economic trends. [2]

• Jasy Liew Suet Yan, Howard R. Turtle, "Exploring Fine-Grained Emotion Detection in Tweets", Proceedings of NAACL-HLT 2016

We examine if common machine learning techniques known to perform well in coarsegrained emotion and sentiment classification can also be applied successfully on a set of fine- grained emotion categories. We first describe the grounded theory approach used to develop a corpus of 5,553 tweets manually annotated with 28 emotion categories. From our preliminary experiments, we have identified two machine learning algorithms that perform well in this emotion classification task and demonstrated that it is feasible to train classifiers to detect 28 emotion categories without a huge drop in performance compared to coarser-grained classification schemes.

Automatic fine-grained emotion detection is a challenging task but we have demonstrated that it is feasible to train a classifier to perform decently well in classifying as many as 28 emotion categories. Our 28 emotion categories is an extension to the six to eight emotion categories commonly-used in the state-ofthe-art (Alm et al., 2005; Aman & Szpakowicz, 2007; Mohammad, 2012). Some of the 28 emotion categories overlap with those found in existing emotion theories such as Plutchik's (1962) 24 categories on the wheel of emotion and Shaver et al.'s (2001) tree-structured list of emotions. Existing emotion theories in psychology are not developed specifically based on emotions expressed in text. Therefore, our emotion categories offer a more fitting framework for the study of emotion in text. [3]

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 4, April 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1004153|

Maryam Hasan, Elke Rundensteiner,Emmanuel Agu, "EMOTEX: Detecting Emotions in Twitter Messages", ASE BIGDATA / SOCIALCOM /CYBERSECURITY Conference, Stanford University, May 27-31, 2014
Social media and microblog tools are increasingly used by individuals to express their feelings and opinions in the form of short text messages. Detecting emotions in text has a wide range of applications including identifying anxiety or depression of individuals and measuring well-being or public mood of a community. In this paper, we propose a new approach for automatically classifying text messages of individuals to infer their emotional states. To model emotional states, we utilize the well-established Circumplex model that characterizes effective experience along two dimensions: valence and arousal. We select Twitter messages as input data set, as they provide a very large, diverse and freely available ensemble of emotions. Using hash-tags as labels, our methodology trains supervised classifiers to detect multiple classes of emotion on potentially huge data sets with no manual effort. We investigate the utility of several features for emotion detection, including unigrams, emoticons, negations and punctuation. To tackle the problem of sparse and high dimensional feature vectors of messages, we utilize a lexicon of emotions. We have compared the accuracy of several machine learning algorithms, including SVM, KNN, Decision Tree, and Naive Bayes for classifying Twitter messages. Our technique has an accuracy of over 90%, while demonstrating robustness across learning algorithms. [4]

- A. Radford, R. Jozefowicz, and I. Sutskever, "Learning to Generate Reviews and Discovering Sentiment", 2017.

Authors explore the properties of byte-level recurrent language models. When given sufficient amounts of capacity, training data, and compute time, the representations learned by these models include disentangled features corresponding to high-level concepts. Specifically, it find a single unit which performs sentiment analysis. These representations, learned in an unsupervised manner, achieve state of the art on the binary subset of the Stanford Sentiment Treebank. They are also very data efficient. When using only a handful of labeled examples, this approach matches the performance of strong baselines trained on full datasets. Authors also demonstrate the sentiment unit has a direct influence on the generative process of the model. Simply fixing its value to be positive or negative generates samples with the corresponding positive or negative sentiment. [5]

- B. Nejat, G. Carenini, and R. Ng, "Exploring Joint Neural Model for Sentence Level Discourse Parsing and Sentiment Analysis", Proc. of the SIGDIAL 2017.

Discourse Parsing and Sentiment Analysis are two fundamental tasks in Natural Language Processing that haqve been shown to be mutually beneficial. In this work, authors design and compare two Neural models for jointly learning both tasks. In this approach, authors first create a vector representation for all the text segments in the input sentence. Next, it apply three different Recursive Neural Net models: one for discourse structure prediction, one for discourse relation prediction and one for sentiment analysis. Finally, authors combine these Neural Nets in two different joint models: Multi-tasking and Pre-training. The results on two standard corpora indicate that both methods result in improvements in each task but Multi-tasking has a bigger impact than Pre-training. Specifically for Discourse Parsing, shows improvements in the prediction on the set of contrastive relations. [6]

- N. Nodarakis, S. Sioutas, A. Tsakalidis, and G. Tzimas, "Using Hadoop for Large Scale Analysis on Twitter: A Technical Report", 2016.

In this paper, authors go one step further and develop a novel method for sentiment learning in the MapReduce framework. their algorithm exploits the hash tags and emoticons inside a tweet, as sentiment labels, and proceeds to a classification procedure of diverse sentiment types in a parallel and distributed manner. Moreover, it utilize Bloom filters to compact the storage size of intermediate data and boost the performance of the algorithm. Through an extensive experimental evaluation, It prove that this solution is efficient, robust and scalable and confirm the quality of sentiment identification. [7]

- Y. Zhang and B. C. Wallace, "A Sensitivity Analysis of (and Practitioners' Guide to) Convolutional Neural Networks for Sentence Classification", 2016.

It is currently unknown how sensitive model performance is to changes in these configurations for the task of sentence classification. Authors thus conduct a sensitivity analysis of one-layer CNNs to explore the effect of architecture components on model performance; the aim is to distinguish between important and comparatively inconsequential design decisions for sentence classification. authors focus on one-layer CNNs (to the exclusion of more complex models) due to their comparative simplicity and strong empirical performance, which makes it a modern standard baseline method akin to Support Vector Machine (SVMs) and logistic regression. Authors derive practical advice from our extensive empirical results for those interested in getting the most out of CNNs for sentence classification in real world settings. [8]

- J. Guo, W. Che, H. Wang, and T. Liu, "A Universal Framework for Inductive Transfer Parsing across Multi-typed Treebanks", 2016.

Various treebanks have been released for dependency parsing. Despite that treebanks may belong to different languages or have different annotation schemes, they contain common syntactic knowledge that is potential to benefit each other. This paper presents a universal framework for transfer parsing across multi-typed treebanks with deep multi-task learning. Authors consider two kinds of treebanks as source: the multilingual universal treebanks and the monolingual heterogeneous treebanks. Knowledge across the source and target treebanks are effectively transferred through multi-level parameter sharing. Experiments on several benchmark datasets in various languages demonstrate that Saif M. Mohammad and Svetlana Kiritchenkothis approach can make effective use of arbitrary source treebanks to improve target parsing models. [9]

- Saif M. Mohammad and Svetlana Kiritchenko, "Using
Hashtags to Capture Fine Emotion Categories from Tweets"

Detecting emotions in microblogs and social media posts has applications for industry, health, and security. Statistical, supervised automatic methods for emotion detection rely on text that is labeled for emotions, but such data is rare and available for only a handful of basic emotions. In this paper, we show that emotionword hashtags are good manual labels of emotions in tweets. We also propose a method to generate a large lexicon of word–emotion associations from this emotionlabeled tweet corpus. This is the first lexicon with realvalued word–emotion association scores. We begin with experiments for six basic emotions and show that the hashtag annotations are consistent and match with the annotations of trained judges. We also show how the extracted tweets corpus and word–emotion associations can be used to improve emotion classification accuracy in a different non-tweets domain. Eminent psychologist, Robert Plutchik, had proposed that emotions have a relationship with personality traits. However, empirical experiments to establish this relationship have been stymied by the lack of comprehensive emotion resources. Since personality may be associated with any of the hundreds of emotions, and since our hashtag approach scales easily to a large number of emotions, we extend our corpus by collecting tweets with hashtags pertaining to 585 fine emotions. Then, for the first time, we present experiments to show that fine emotion categories such as that of excitement, guilt, yearning, and admiration are useful in automatically detecting personality from text. Streamof-consciousness essays and collections of Twitter posts marked with personality traits of the author are used as the test sets. [10]

III. PROPOSED METHODOLOGY

As the most popular micro blogging platform, Twitter has a vast amount of information available in the form of tweets shared by millions of users. Since this data stream is constantly growing, it is difficult to extract relevant information for users. More and more people want to benefit from these data and get a personalized service from Twitter. Extracting the semantic meaning of Twitter and modeling the interests of users allows people to enjoy a personalized service on Twitter. First it evaluate and compare the performance of proposed approaches of emotion detection. Then use these approaches of emotion detection to analyze Twitter sample dataset for the purpose of user modeling.

A. Modules

- Textual Context Recognition

  The process involves collecting raw data related to feedback and pre proccing it. After classification of prprocessed data. Feedback are classified according to sentiments such as Happy, sad, angry, surprised, etc,

- Face Detection

  Given a picture, detecting the presence of a person's face may be a complex task thanks to the possible variations of the face. The various sizes, angles and poses human face may need within the image cause this variation. The emotions which are deducible from the face and different like conditions such as illumination and occlusions also affect facial appearances. The approaches of the past few decades in face detection are often classified into four: knowledge-based approach, feature invariant approach, template –based approach and appearance-based approach.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 4, April 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1004153|

- Facial Feature Extraction

  Contracting the facial muscles produces changes in both the direction and magnitude of skin surface displacement, and within the appearance of permanent and transient countenance. Samples of permanent features are eyes, brow, and any furrows that become permanent with age. Transient features include facial lines and furrows that aren't present at rest. So as to research a sequence of images, we assume that the primary frame may be a neutral expression. After initializing the templates of the permanent features within the first frame, both geometric countenance and Gabor wavelets coefficients are automatically extracted

B. Algorithms

1) Algorithm :Support Vector Machine: In data analytics or decision sciences most of the time I come across the situations where I need to classify our data based on a certain dependent variable. To support the solution for this need there are multiple techniques which can be applied; Logistic Regression, Random Forest Algorithm, Bayesian Algorithm are a few to name. SVM is a machine learning technique to separate data which tries to maximize the gap between the categories.

Algorithm for classification of emotion.
Input: D Dataset, Semantic of Tokens, Tweets;
Output: Classification of Application
Step 1: for each tweet tweet id in D do Step 2: Get on-demand features and stored on vector x for tweet id
Step 3: x.add (Get Features(tweet id));
Step 4: end for
Step 5: for each tweet in x vector do
Step 6: Fetch first feature and stored in b, and other features in w
Step 7: h w, b (x) = g (z) here, z = (w T x + b)
Step 8: if (z ≤ 0)
Step 9: assign g(z) = 1; Step 10: else g(z) = -1;
Step 11: end if

2) Algorithm : Convolutional Neural Network (CNN): A Convolutions Neural Network (CNN) is comprised of one or more convolutions layers (often with a sub sampling step) and then followed by one or more fully connected layers as in a standard multi layer neural network. The architecture of a CNN is designed to take advantage of the 2D structure of an input image (or other 2D input such as a speech signal). This is achieved with local connections and tied weights followed by some form of pooling which results in translation in variant features. Another benefit of CNN's is that they are easier to
train and have many fewer parameters than fully connected networks with the same number of hidden units.

- Step 1: Dataset containing images along with reference caption is fed into the system
- Step 2: The convolutional neural network is used a encoder which extracts image features 'f' pixel by pixel.
- Step 3: Matrix factorization is performed on the extracted pixels. The matrix is of m x n.
- Step 4: Max pooling is performed on this matrix where maximum value is selected and again fixed into matrix.
- Step 5: Normalization is performed where the every negative value is converted to zero.
- Step 6: To convert values to zero rectified linear units are used where each value is filtered and negative value is Fig. 3. Screenshot 3 set to zero.
- Step 7: The hidden layers take the input values from the visible layers and assign the weights after calculating maximum probability.

Step 8: assign g(z) = 1;
Step 10: else g(z) = -1;
Step 11: end if

2) Algorithm : Convolutional Neural Network (CNN): A Convolutions Neural Network (CNN) is comprised of one or more convolutions layers (often with a sub sampling step) and then followed by one or more fully connected layers as in a standard multi layer neural network. The architecture of a CNN is designed to take advantage of the 2D structure of an input image (or other 2D input such as a speech signal). This is achieved with local connections and tied weights followed by some form of pooling which results in translation in variant features. Another benefit of CNN's is that they are easier to modeling on the Twitter platform and seeks to combine two hotspots.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 4, April 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1004153|

IV.RESULTS & DISCUSSIONS

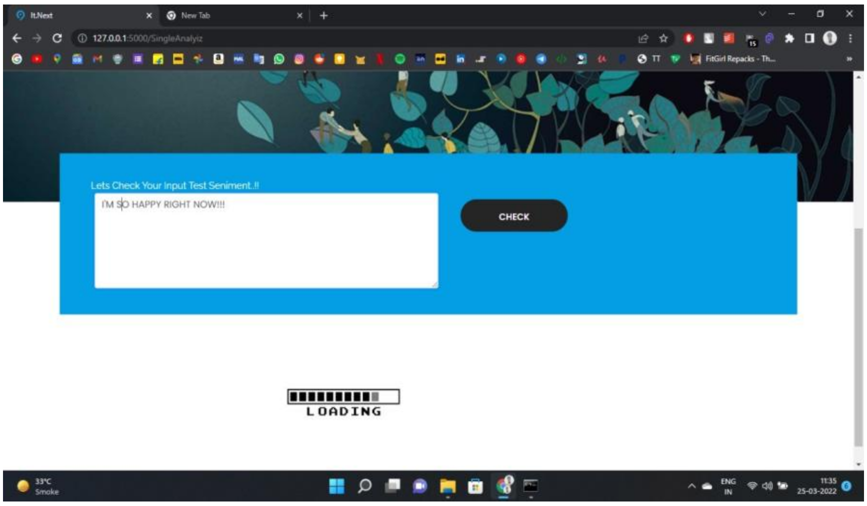


Fig. 1. Screenshot 1 recognize in tweeter

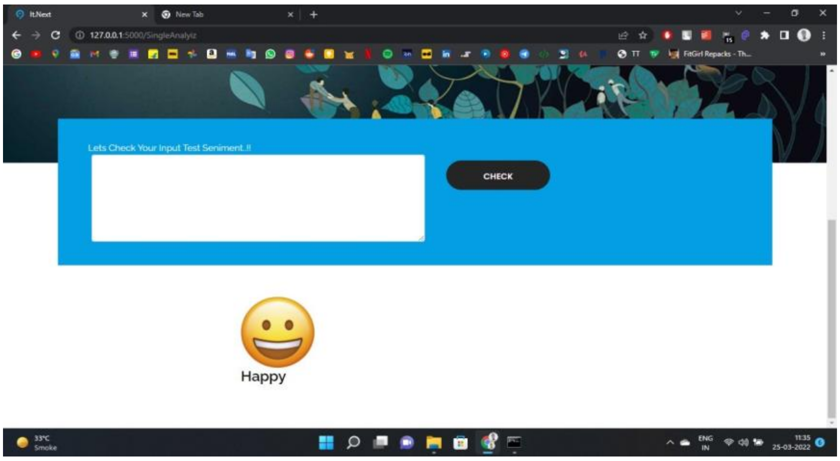


Fig. 2. Screenshot 2

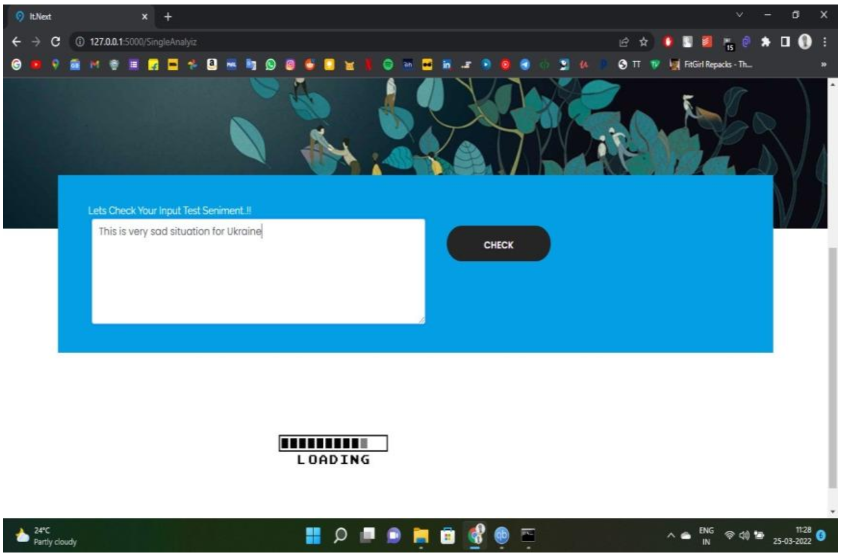


Fig. 3. Screenshot 3

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 4, April 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1004153|

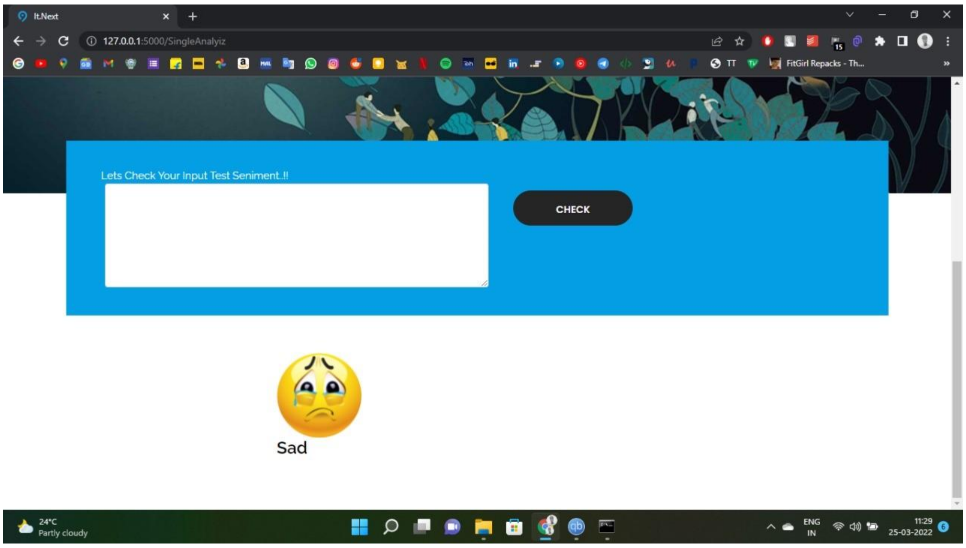


Fig.4.Screenshot4

V.CONCLUSION

The main idea of the paper is to expand the leverage of deep learning and data mining techniques for emotion. Proposed system is directly connected to the tweeter's public account to get the tweets stream. Experiment shows the result in three algorithms name as SVM, Random Forest and CNN. In experiments needs to tokenize the tweets and get semantic of the token to get better result. Accuracy of the proposed system using semantic with preprocessing is much better than existing system. Proposed system is taken word as input instead of characters.This proposed System works on probably the largest data set for emotion prediction, using tweets from years. Since the training data was annotated automatically and since I use character-based approaches, our solution is language independent and could easily be adapted for other languages. I believe this work is beneficial for the user modeling on the Twitter platform and seeks to combine two hotspots.

REFERENCES

[1] Niko Colneric and Janez Demsar, "Emotion Recognition on Twitter: Comparative Study and Training a Unison Model", IEEE TRANSACTIONS ON AFFECTIVE COMPUTING, FEBRUARY 2018.

[2] A. Radford, R. Jozefowicz, and I. Sutskever, "Learning to Generate Reviews and Discovering Sentiment", 2017.

[3] B. Nejat, G. Carenini, and R. Ng, "Exploring Joint Neural Model for Sentence Level Discourse Parsing and Sentiment Analysis", Proc. of the SIGDIAL 2017.

[4] N. Nodarakis, S. Sioutas, A. Tsakalidis, and G. Tzimas, "Using Hadoop for Large Scale Analysis on Twitter: A Technical Report", arXiv preprint arXiv: 1602.01248, 2016.

bibitemb4 Y. Zhang and B. C. Wallace, "A Sensitivity Analysis of (and Practitioners' Guide to) Convolutional Neural Networks for Sentence Classification", ArXiv preprint arXiv:1510.03820v4, 2016.

[5] J. Guo, W. Che, H. Wang, and T. Liu, "A Universal Framework for Inductive Transfer Parsing across Multi-typed Treebanks", Proc. of the 26th Int. Conf. on Computational Linguistics (COLING-16), pp. 12–22, 2016.

[6] S. M. Mohammad and S. Kiritchenko, "Using Hashtags to Capture Fine Emotion Categories from Tweets", Computational Intelligence, vol. 31, no. 2, pp. 301–326, 2015.

[7] B. Plank and D. Hovy, "Personality Traits on Twitter or How to Get
1,500 Personality Tests in a Week", in Proc. of the 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, 2015.

[8] D. Bamman and N. A. Smith, "Contextualized Sarcasm Detection on Twitter", in Proc. of the 9th Int. AAAI Conf. on Web and Social Media. Citeseer, 2015.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 4, April 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1004153|

[9] X. Liu, J. Gao, X. He, L. Deng, K. Duh, and Y. Y. Wang, "Representation Learning Using Multi-Task Deep Neural Networks for Semantic Classification and Information Retrieval", Proc.of the Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 921, 2015.

[10] F. Feng, X. Wang, and R. Li, "Cross-modal Retrieval with Correspondence Autoencoder", Proc.of the 22nd ACM Int. Conf. on Multimedia, pp. 7–16, 2014.

[11] Zhaolong Li, "Analyzing Emotion on Twitter for User Modeling", Master's Thesis, October 2013.

[12] K. Gimpel, N. Schneider, B. O'Connor, D. Das, D. Mills, J. Eisenstein, M. Heilman, D. Yogatama, J. Flanigan, and N. A. Smith, "Part-of-
Speech Tagging for Twitter: Annotation, Features, and Experiments," in
Proc. of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, vol. 2, no. 2. ACL, 2011, pp. 42–47.

[13] I. T. Jolliffe, "Principal Component Analysis," in Springer Series in Statistics, 2002.

[14] R. B. Bradford, "An Empirical Study of Required Dimensionality for Large-scale Latent Semantic Indexing Applications", Proc.of the 17th ACM Conf., p. 153, 2008.

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.165

doi
crossref

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com