International Journal of Early Childhood Special Education (INT-JECSE)
DOI:10.48047/INTJECSE/V15I5.16 ISSN: 1308-5581 Vol 15, Issue 05 2023

EMOTION DETECTION FROM MICRO-BLOGS USING NOVEL INPUT REPRESENTATION

[1] Dr. S Sowjanya, [2] Namshior Karthik,[3] Jarapala Srikanth Naik,[4] Kethavath Divakar ,[5] *Koppolu Venu*
[1] Asst Professor, Department of CSE, Sreyas Institute of Engineering and Technology, Telangana, India.
[1] soujikits@gmail.com
[2,3,4,5] Department of CSE, Sreyas Institute of Engineering and Technology, Telangana, India
[2] namshiork4@gmail.com , [3] ***srikanthjarapala701@gmail.com*** , [4] kethavathdivakar2002@gmail.com , [5] kvenuyadav1234@gmail.com

Abstract
Human behavior, social interaction, and decision-making are all influenced by emotion, which is a basic, intrinsic state of mind. Owing to the Internet's explosive growth in the modern era, online social media (OSM) platforms are now widely used as a means of opinion and emotion expression. As a result of the development of artificial intelligence (AI) algorithms-driven natural language processing (NLP) methods, emotion detection (ED) from user-generated OSM data has become a heavily researched area. The brief, informal, and unstructured texts that are typical on micro blogging sites like Twitter make it difficult to extract useful features for spotting observable patterns. In this paper, we present a novel feature representation that can capture users' emotional states, taken from user-generated Twitter data. An enhanced method predicated on RF, The input representation is built using SVM, KNN, XGBOOST, and other techniques. It is made up of linguistic, sentiment, and stylistic elements that are taken from tweets. Using the new feature representation, a voting ensemble classifier with algorithmically optimized weights is presented to improve the accuracy of emotion detection. A benchmark Twitter emotion detection dataset is used to train and evaluate the suggested classifier. Each sample in the dataset is labelled with one of the six classes: fear, surprise, anger, sadness, joy, and love. The experimental results show that the suggested method achieves the highest accuracy and outperforms the most advanced classical machine learning-based emotion detection techniques.
Keywords: RF, SVM, KNN, XGBOOST.

I INTRODUCTION

As technology continues to advance, the proliferation of user-generated content on online social media (OSM) platforms has made opinion mining an important domain of research. Nowadays, individuals are increasingly influenced by the innovative features and trends introduced by different OSM platforms. Over the last few years, these factors resulted in a dramatic increase in users' interest in different social media platforms.

According to the Digital 2022 Global StatshotReport, there were a staggering 4.62 billion active social media users globally in January 2022, a 10.10% increase from the previous year. Online social media platforms have become such a dominant force that, along with traditional face-to-face and electronic media communication, people's interactions are profoundly shaped by sharing their opinions, thoughts, and stories about global events online. In addition, these platforms provide access to a plethora of user-generated data that are leveraged by practitioners across various sectors for effective decision-making in business and technological intervention.

Online social media provides users with platforms and opportunities to express, communicate, and share their opinions, views, and thoughts. The user-generated social media content reveals valuable insights into people's emotional states that can inform a wide range of behavioral and psychological stances of an individual. Emotion is essential in every aspect of a person's life, influencing decision-making, social relationships, and behavior. As a result, there has been a surge in research on the application of artificial intelligence (AI) and machine learning (ML) techniques to infer users' emotional states from their social media content . Automatic emotion detection (ED) involves using natural language processing (NLP) techniques and machine learning algorithms to decipher an individual's emotional states . In online social media platforms, people share content in various data formats such as text, image, video, graphics interchange format (GIF), and others that inform users' emotional, psychological, and personality traits. While emotion detection from facial expression, and more recently gait have been widely explored, detecting emotions from textual data is still an emerging research area.

II LITERATURESURVEY

AI based emotion detection for textual big data: Techniques and contribution:

Online Social Media (OSM) like Facebook and Twitter has emerged as a powerful tool to express via text people's opinions and feelings about the current surrounding events. Understanding the emotions at the fine-grained level of these expressed thoughts is important for system improvement. Such crucial insights cannot be completely obtained by doing AI-based big data sentiment analysis; hence, text-based emotion detection using AI in social media big data has become an upcoming area of Natural Language Processing research. It can be used in various fields such as understanding expressed emotions, human–computer interaction, data mining, online education, recommendation systems, and psychology. Even though the research work is ongoing in this domain, it still lacks a formal study that can give a qualitative (techniques used) and quantitative (contributions) literature overview. This study has considered 827 Scopus and 83 Web of Science research papers from the years 2005–2020 for the analysis. The qualitative review represents different emotion models, datasets, algorithms, and application domains of text-based emotion detection. The quantitative bibliometric review of contributions presents research details such as publications, volume, co-authorship networks, citation analysis, and demographic research distribution. In the end, challenges and probable solutions are showcased, which can provide future research directions in this area.

Biases, fairness, and implications of using AI in social media data mining:

Online social media (OSM) has become an integral part of an individual's daily life. The extensive computational power and decision-making ability of artificial intelligence (AI) and the proliferation of user-generated data on OSM have made the opinion one of the key emerging research areas. However, the ease of accessing, manipulating, and mining such user-generated data raises concerns about privacy and security, data and algorithmic biases and fairness. Nevertheless, their personal and societal implications are barely addressed. In this paper, we discuss the limitations, fairness, and biases introduced in data mining and the AI model development. Moreover, we describe the possible implications of using AI systems on users' privacy and address future research directions to mitigate potential biases.

The human in emotion recognition on social media: Attitudes, outcomes, risks:

Emotion recognition algorithms recognize, infer, and harvest emotions using data sources such as social media behavior, streaming service use, voice, facial expressions, and biometrics in ways often opaque to the people providing these data. People's attitudes towards emotion recognition and the harms and outcomes they associate with it are important yet unknown. Focusing on social media, we interviewed 13 adult U.S. social media users to fill this gap. We find that people view emotions as insights to behavior, prone to manipulation, intimate, vulnerable, and complex. Many find emotion recognition invasive and scary, associating it with autonomy and control loss. We identify two categories of emotion recognition's risks: individual and societal. We discuss findings' implications for algorithmic accountability and argue for considering emotion data as sensitive. Using a Science and Technology Studies lens, we advocate that technology users should be considered as a relevant social group in emotion recognition advancements.

Deep facial expression recognition: A survey:

With the transition of facial expression recognition (FER) from laboratory-controlled to challenging in-the-wild conditions and the recent success of deep learning techniques in various fields, deep neural networks have increasingly been leveraged to learn discriminative representations for automatic FER. Recent deep FER systems generally focus on two important issues: overfitting caused by a lack of sufficient training data and expression-unrelated variations, such as illumination, head pose and identity bias. In this paper, we provide a comprehensive survey on deep FER, including datasets and algorithms that provide insights into these intrinsic problems. First, we introduce the available datasets that are widely used in the literature and provide accepted data selection and evaluation principles for these datasets. We then describe the standard pipeline of a deep FER system with the related background knowledge and suggestions of applicable implementations for each stage. For the state of the art in deep FER, we review existing novel deep neural networks and related training strategies that are designed for FER based on both static images and dynamic image sequences, and discuss their advantages and limitations. Competitive performances on widely used benchmarks are also summarized in this section. We then extend our survey to additional related issues and application scenarios. Finally, we review the remaining challenges and corresponding opportunities in this field as well as future directions for the design of robust deep FER systems.

Facial metamorphosis using geometrical methods for biometric applications:

Facial expression modeling has been a popular topic in biometrics for many years. One of the emerging recent trends is capturing subtle details such as wrinkles, creases and minor imperfections that are highly important for biometric modeling as well as matching. In this paper, we suggest a novel approach to the problem of expression modeling and morphing based on a geometry-based paradigm. In 2D image space, a distance-based morphing system is utilized to create a line drawing style facial animation from two input images representing frontal and profile views of the face. Aging wrinkles and expression lines are extracted and mapped back to the synthesized facial NPR (nonphotorealistic) sketches. In 3D object space, we present a metamorphosis system that combines the traditional free-form deformation (FFD) model with data interpolation techniques based on the proximity preserving Voronoi diagram. With feature points selected from two images of the target face, the proposed system generates the 3D target facial model by transforming a generic model. Experimental results demonstrate that morphing sequences generated by our systems are of convincing quality

III EXISTINGSYSTEM

In text-based emotion detection, natural language processing (NLP) techniques are employed to extract meaningful patterns from the text data, that are leveraged by ML algorithms to infer the users' emotions. Text-based emotion detection has important applications in various fields, including customer service, mental health support, opinion mining, and personalization . It enables machines to recognize and understand the emotions expressed in text, which can improve communication and has the potential to enhance user experience. However, automatic emotion detection from OSM textual data is challenging due to the scarcity of publicly available datasets labeled with emotion categories, the unstructured nature of the user-generated data, and the need for user's privacy . Moreover, emotion recognition is especially critical for user-generated tweets. They consist of short, informal, and unstructured text; incomplete, misspelled, and slang words; abbreviations, acronyms, and special characters, all of which require extensive text preprocessing and accurate feature extraction

DISADVANTAGES:

However, automatic emotion detection from OSM textual data is challenging due to the scarcity of publicly available datasets labeled with emotion categories, the unstructured nature of the user-generated data, and the need for user's privacy

Moreover, emotion recognition is especially critical for user-generated tweets. They consist of short, informal, and unstructured text; incomplete, misspelled, and slang words; abbreviations, acronyms, and special characters, all of which require extensive text preprocessing and accurate feature extraction.

IV PROBLEM STATEMENT

Human behavior, social interaction, and decision-making are all influenced by emotion, which is a basic, intrinsic state of mind. Owing to the Internet's explosive growth in the modern era, online social media (OSM) platforms are now widely used as a means of opinion and emotion expression. The brief, informal, and unstructured texts that are typical on micro blogging sites like Twitter make it difficult to extract useful features for spotting observable patterns. In this paper, we present a novel feature representation that can capture users' emotional states, taken from user-generated Twitter data

V PROPOSED SYSTEM

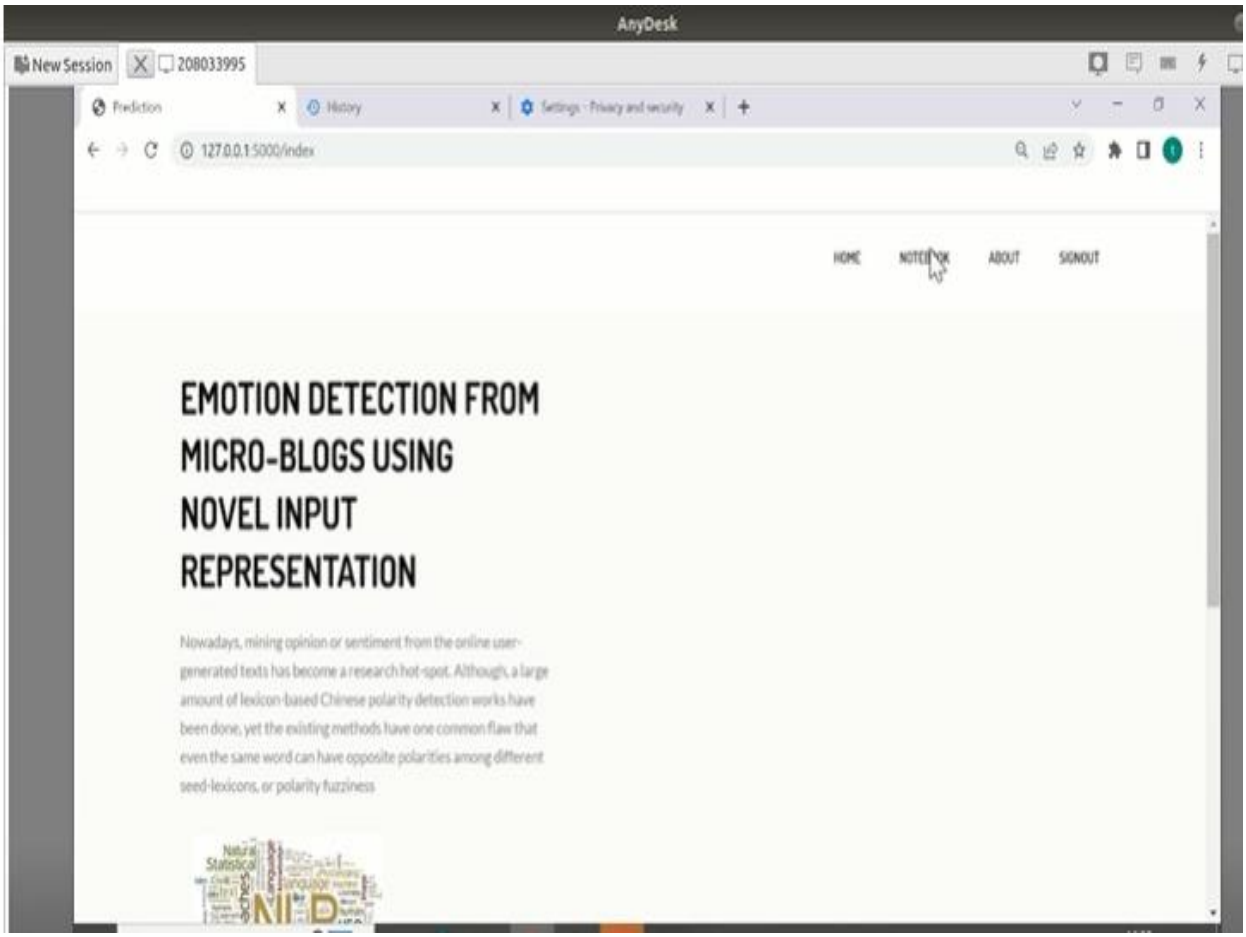
In this paper, we introduce a novel representation of features extracted from user-generated Twitter data that can capture users' emotional states. An advanced approach based on RF, SVM, KNN, XGBOOST and more are used to construct the input representation which is composed of stylistic, sentiment, and linguistic features extracted from tweets. A voting ensemble classifier with weights optimized by a algorithms are introduced to increase the accuracy of emotion detection using the novel feature representation. The proposed classifier is trained and tested on a benchmark Twitter emotion detection dataset where each sample is labeled with either of the six classes: sadness, joy, love, anger, fear, and surprise.

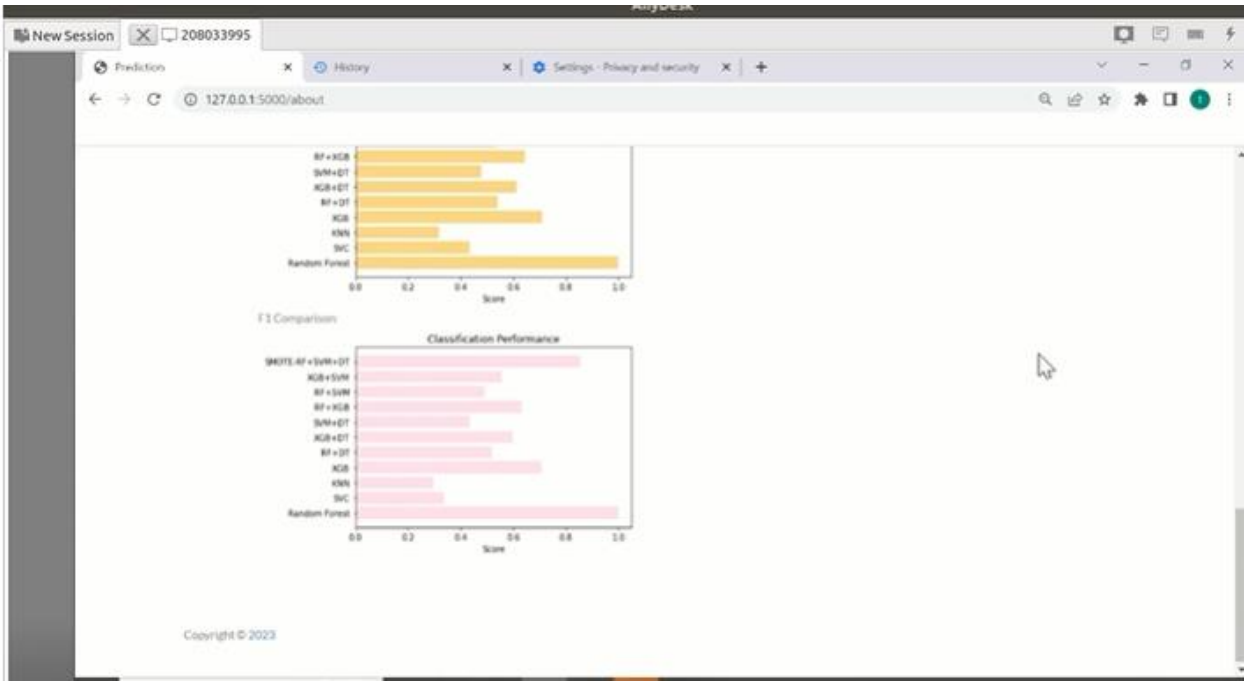
VI IMPLEMENTATION

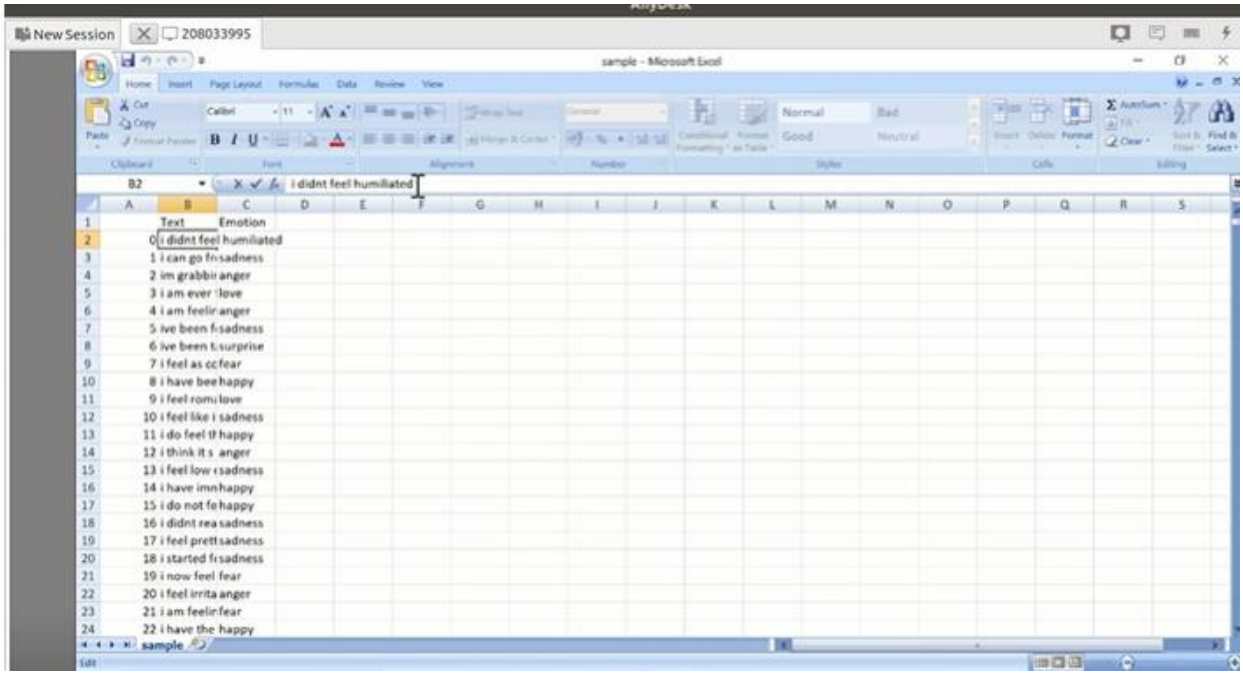
- Data exploration: using this module we will load data into system
- Processing: Using the module we will read data for processing
- Splitting data into train & test: using this module data will be divided into train & test
- Model generation: Algorithms accuracy calculated

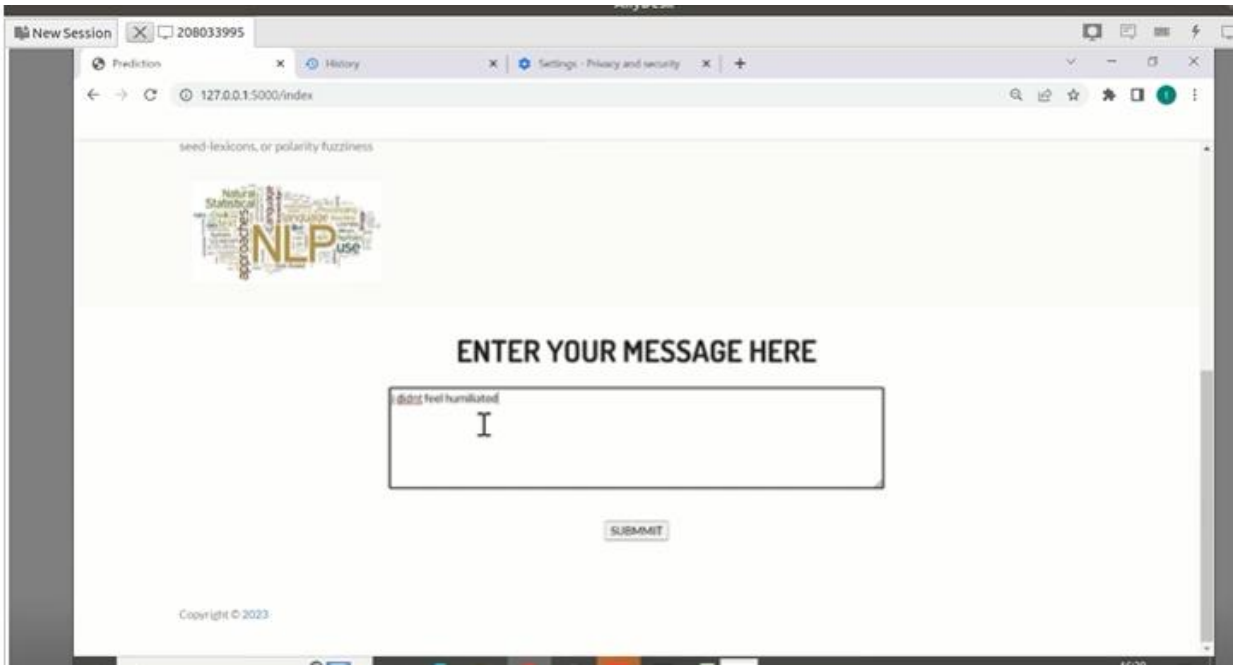
- User signup & login: Using this module will get registration and login
- User input: Using this module will give input for prediction

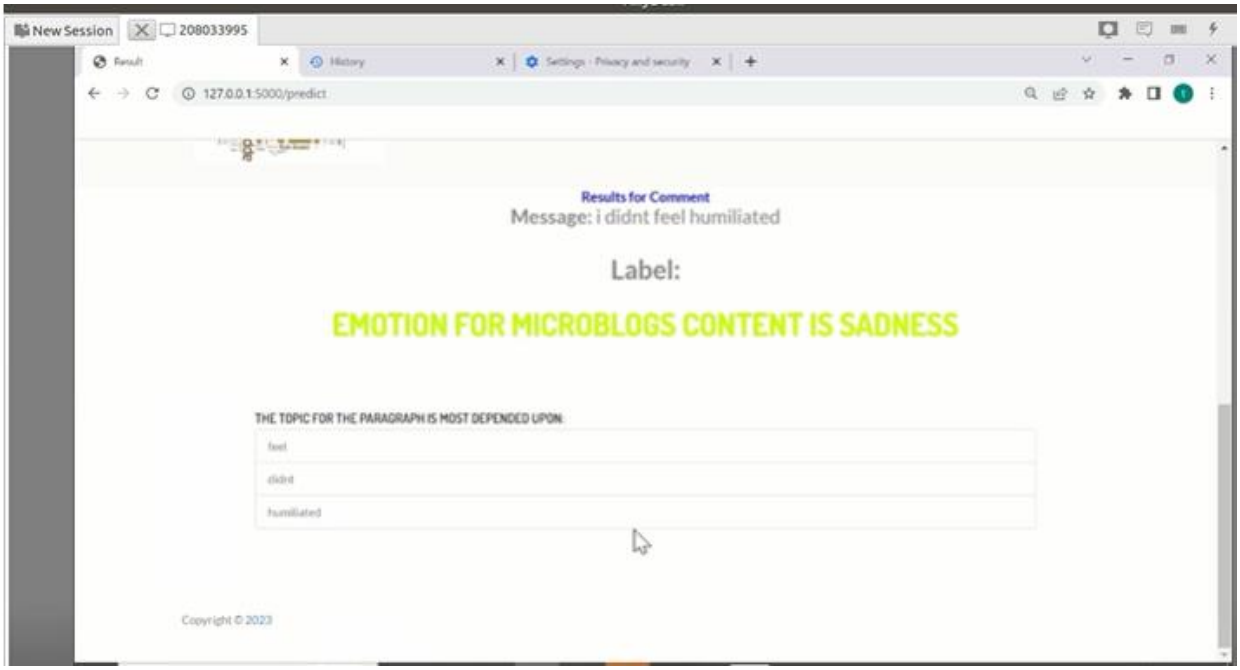
VII RESULTS




New Session
208033995
127.0.0.1:5000/about
Classification Performance
Score
Copyright © 2023


New Session
208033995
sample - Microsoft Excel
i didnt feel humiliated
Text
Emotion
sample





VIII CONCLUSION

This paper proposed a novel input representation SSEL by combining the stylistic (S), sentiment (SE), and linguistic (L) features extracted from tweets for representing users' emotional states. A genetic algorithm was leveraged to combine and compress the distinct feature sets into one unified representation. This paper also presented a novel combination of linear support vector machine, XGBoost, and random forest as a weighted average voting classifier for detecting emotions by classifying the tweets into six independent categories using the proposed input representation. This research showed that the stylistic and sentiment attributes when combined with the language-based input representation can capture discernible patterns in the tweets that are highly predictive for emotion

detection. The proposed emotion detection approach was compared with five independent classical ML classifiers, six different combinations of weighted ensemble voting classifiers, and four recent state-of-the-art ML-based ED techniques by employing the proposed input representations extracted from a publicly available Twitter emotion detection dataset. The experimental results show that our proposed ED system outperforms all the recent approaches considering each of the performance evaluation metrics and establishes a new performance benchmark for the experimented dataset

REFERENCES

[1] P. A. Rauschnabel, P. Sheldon, and E. Herzfeldt, ‘‘What motivates users to hashtag on social media?’’ Psychol. Marketing, vol. 36, no. 5, pp. 473–488, May 2019.

[2] S. Kemp. (2022). Digital 2022: Digital Adoption Doubled Over the Past Decade. [Online]. Available: https://datareportal.com/reports/digital- 2022-digital-adoption-doubled-over-the-past-decade

[3] S. Kusal, S. Patil, K. Kotecha, R. Aluvalu, and V. Varadarajan, ‘‘AI based emotion detection for textual big data: Techniques and contribution,’’ Big Data Cognit. Comput., vol. 5, no. 3, p. 43, Sep. 2021.

[4] M. L. Gavrilova, F. Anzum, A. Hossain Bari, Y. Bhatia, F. Iffath, Q. Ohi, M. Shopon, and Z. Wahid, ‘‘A multifaceted role of biometrics in online security, privacy, and trustworthy decision making,’’ in Breakthroughs in Digital Biometrics and Forensics. Cham, Switzerland: Springer, 2022, pp. 303–324.

[5] F. Anzum, A. Z. Asha, and M. L. Gavrilova, ‘‘Biases, fairness, and implications of using AI in social media data mining,’’ in Proc. Int. Conf. Cyberworlds (CW), Sep. 2022, pp. 251–254.

[6] N. Andalibi and J. Buss, ‘‘The human in emotion recognition on social media: Attitudes, outcomes, risks,’’ in Proc. CHI Conf. Hum. Factors Comput. Syst., Apr. 2020, pp. 1–16.

[7] S. Li and W. Deng, ‘‘Deep facial expression recognition: A survey,’’ IEEE Trans. Affect. Comput., vol. 13, no. 3, pp. 1195–1215, Jul. 2022.

[8] Y. Luo, M. L. Gavrilova, and P. S. P. Wang, ‘‘Facial metamorphosis using geometrical methods for biometric applications,’’ Int. J. Pattern Recognit. Artif. Intell., vol. 22, no. 3, pp. 555–584, May 2008.

[9] Y. Bhatia, A. H. Bari, and M. Gavrilova, ‘‘A LSTM-based approach for gait emotion recognition,’’ in Proc. IEEE 20th Int. Conf. Cognit. Informat. Cognit. Comput. (ICCI*CC), Oct. 2021, pp. 214–221.

[10] P. Nandwani and R. Verma, ‘‘A review on sentiment analysis and emotion detection from text,’’ Social Netw. Anal. Mining, vol. 11, no. 1, pp. 1–19, Dec. 2021.

[11] Y. HaCohen-Kerner, D. Miller, and Y. Yigal, ‘‘The influence of preprocessing on text classification using a bag-of-words representation,’’ PLoS ONE, vol. 15, no. 5, May 2020, Art. no. e0232525.

[12] J. Fürnkranz, ‘‘A study using n-gram features for text categorization,’’ Austrian Res. Inst. Artif. Intell., vol. 3, no. 1998, pp. 1–10, 1998.

[13] S. Qaiser and R. Ali, ‘‘Text mining: Use of TF-IDF to examine the relevance of words to documents,’’ Int. J. Comput. Appl., vol. 181, no. 1, pp. 25–29, Jul. 2018.

[14] K. W. Church, ‘‘Word2Vec,’’ Nat. Lang. Eng., vol. 23, no. 1, pp. 155–162, 2017.

[15] J. Pennington, R. Socher, and C. Manning, ‘‘GloVe: Global vectors for word representation,’’ in Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP), 2014, pp. 1532–1543