EMOTIOM RECOGNITION USING TEXTUAL TWEETS

PROJECT PHASE - 1

TEAM MEMBERS

YUVARAJ P (111720102179)
VENKATESH T (111720102170)
YADESH HARIHAR S (111720102314)

SUPERVISOR
Mr. M.P KARTHIKEYAN, M.Tech., (Ph.D).,
Assistant Professor
CSE

ABSTRACT

- With the exponential growth of social media, especially text data. Sentiment analysis is a technique used to analyze different people's attitudes, emotions, and opinions about anything and can be performed on tweets to analyze public opinion about news, policies, social movements, and personalities. Twitter is used to gather opinions about products, trends and politics. Using advanced machine learning techniques, the proposed model aims to accurately identify and categorize various emotions expressed in tweets, such as happiness, sadness, anger, fear, and others. Utilizing advanced machine learning algorithms and sentiment analysis tools, the model demonstrates robustness in detecting and categorizing complex emotional nuances expressed in short text messages. Through extensive experimentation and validation, the effectiveness of the proposed approach is evaluated and its potential applications in various fields such as social media analysis, mental health assessment, and customer sentiment analysis are highlighted, contributing to a more accurate understanding of human emotions in the digital era.

Key words: Sentiment Analysis, Natural Language Processing, Emotion Detection, Textual Emotion Recognition.

SCOPE OF THE PROJECT

The scope of emotion recognition using textual tweets encompasses the development of advanced natural language processing (NLP) algorithms capable of accurately identifying and categorizing a wide range of emotions expressed in text.
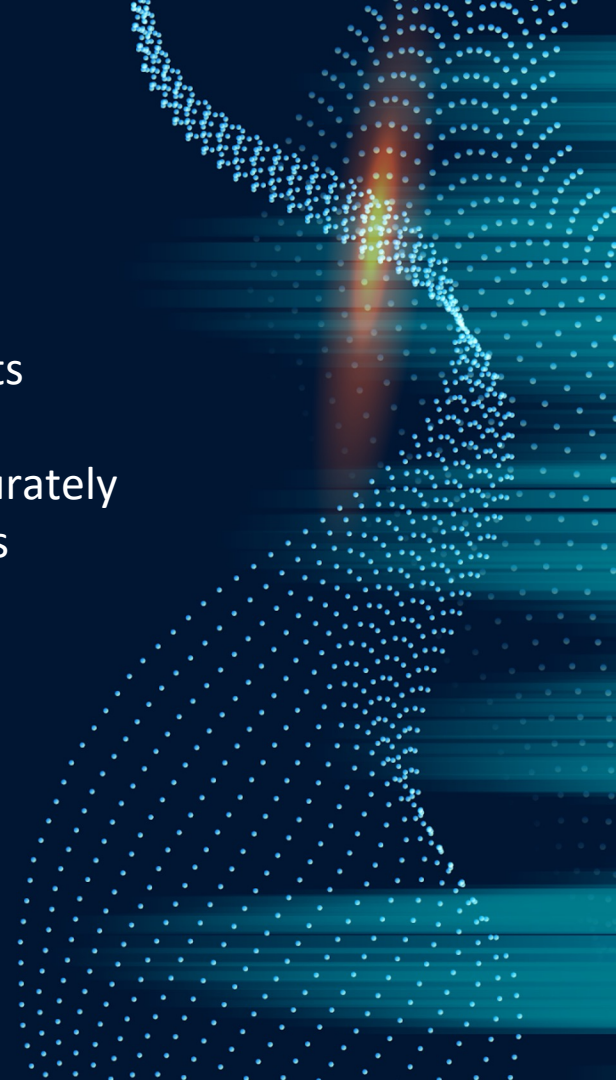
PROBLEM STATEMENT

Understanding and accurately identifying emotions from text tweets presents challenges due to the inherent complexity of human expression in written form. Ambiguities in language, including sarcasm, irony, and slang, present difficulties in correctly deciphering the underlying sentiment. In addition, cultural and contextual nuances further complicate the process, which can lead to misinterpretations. Moreover, the subjective nature of emotions and the lack of explicit indicators often hinder the development of robust and universally applicable models of emotion recognition.
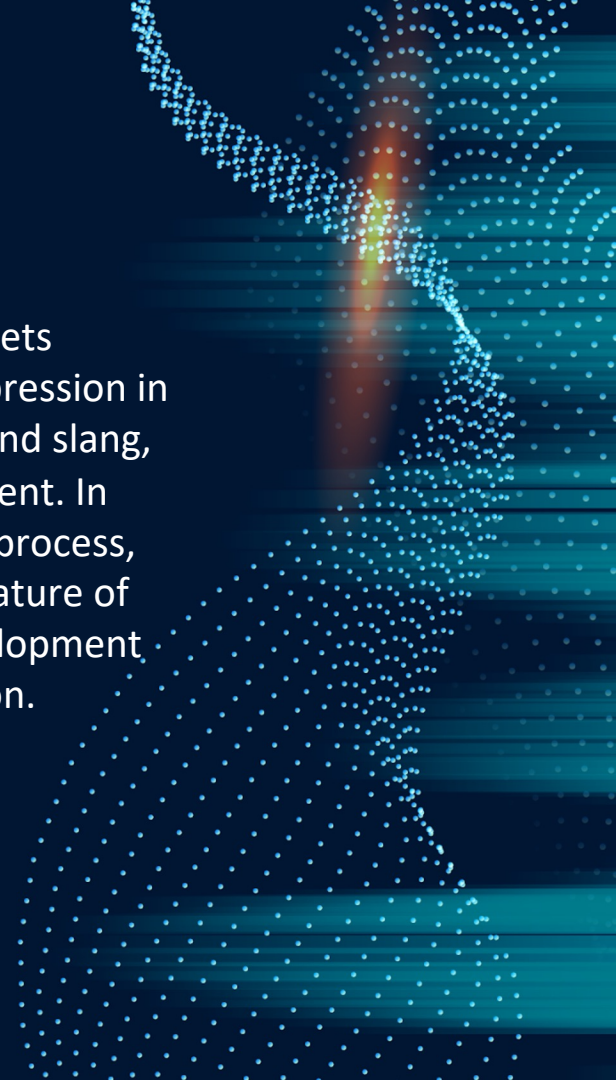
LITERATURE SURVEY

The research paper titled "Twitter sentiment analysis" authored by C. Kariya and P. Khodke, The research is divided into two main phases. In the first phase, researchers conducted a case study to explore existing theories and the most frequently used theories in the current context. This phase will involve reviewing academic materials, articles, and other resources to understand the various methods used in analytical thinking. In the second stage, researchers define the necessity and functionality of the application. This means that they describe the needs and specifications for the evaluation of emotions, as well as the activities and processes involved in their development. This stage may refer to the application of the emotional processes learned in the first stage. In general, research involves investigating scientific theories in existing literature and then using this information to develop methods to voluntarily identify customer needs expressed in tweets and classify them as positive or negative. [1]
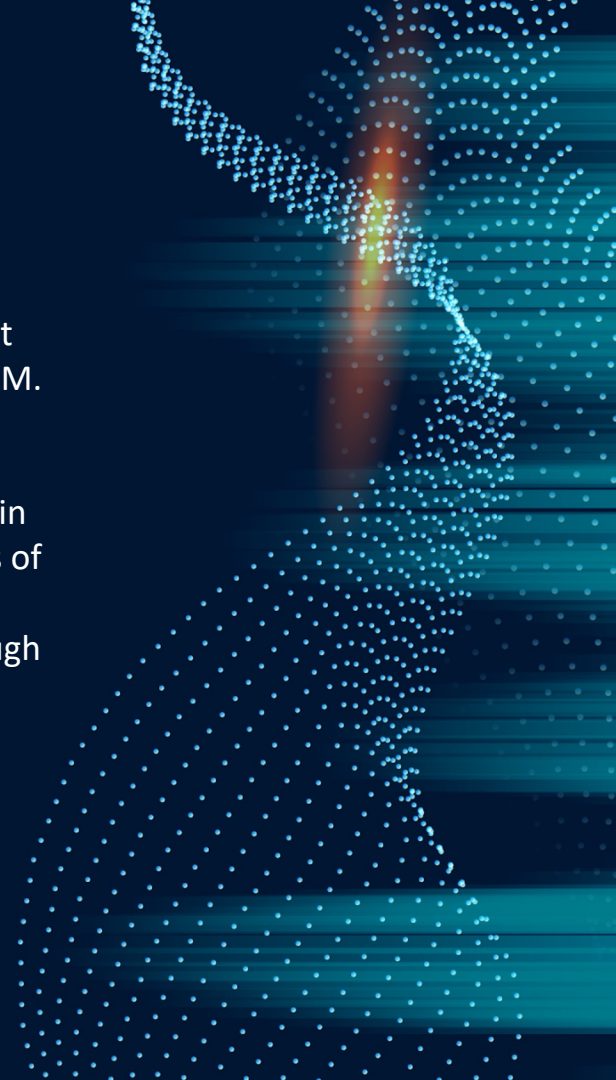
LITERATURE SURVEY

The research paper titled "A study on sentiment analysis techniques of Twitter data" authored by A. Alsaeedi and M. Zubair, This method involves training learning models that use Twitter profile names to predict tweet sentiment. The model learns from its knowledge base to classify new tweets as positive, negative, or neutral. Community Sentiment Analysis: Community Sentiment Analysis combines multiple models to increase accuracy. In this case, the authors will need to combine various analysis theories to create a more reliable and accurate analysis of sentiment on Twitter profiles. The authors compare the effectiveness and results of these different approaches. They aim to determine which method is most effective when applied to Twitter data by evaluating the accuracy and effectiveness of each method. This comparative analysis will provide a better understanding of the advantages and limitations of different theoretical frameworks and guide researchers and practitioners in selecting appropriate models for analysis.[2]

LITERATURE SURVEY

The research paper titled "Twitter benchmark dataset for arabic sentiment analysis" authored by D. Gamal, M. Alfonse, E.-S. M. El-Horbaty, and A.-B. M. Salem, built Twitter benchmark dataset for Arabic Sentiment Analysis. A benchmark Arabic dataset suggested in experiment for estimation investigation demonstrating social event strategy about the latest tweets in various Arabic vernaculars. The experiment dataset incorporates in excess of 151,000 unique assessments which marked into two classes, negative and positive. ML algorithms are functioned in SC; ML algorithm attached through learning arrangements. Sentiment analysis ordinarily executed using one fundamental methodology from a ML(lexicon-based approach) based approach. The calculations functioned via SC on the dataset accomplished 99.90% precision utilizing TF-IDF. [3]

LITERATURE SURVEY

The research paper titled "Fuzzy rule based unsupervised sentiment analysis from social media posts" authored by Srishti Vashishtha, Seba Susan, developed the sentiment analysis of social media posts based on a set of fuzzy rules, which involved multiple lexicons and datasets. The fuzzy system integrated the Natural Language Processing (NLP) techniques and Word Sense Dis- ambiguation based on a novel unsupervised nine fuzzy rule-based system for classification. The scheme provided the exact sentiment values and also dealt with the linguistic problems. Experimentation result showed that the scheme achieved higher performance as compared to the other state-of-the-art methods. However, the scheme exhibits a high error rate, which leads to inaccurate classification. [4]

LITERATURE SURVEY

The research paper titled "Social media opinion summarization using emotion cognition and convolutional neural networks" authored by Peng Wu, Xiaotong Li, Si Shen, Daqing He, developed an Ortony-Clore-Collins (OCC) model and a Convolutional Neural Network (CNN) based opinion summarization method for Chinese microblogging systems. The scheme highlighted the potential of combined emotion cognition with deep learning in sentiment analysis of social media data. The experimental result demonstrated that the scheme outperformed the other state-of-the-art methods in terms of classification and recognition performance. But, the scheme had a complexity of microblog sentiment classification. [5]

PROPOSED SYSTEM

The proposed system for emotion recognition using text tweets integrates advanced techniques, including deep learning architectures and transformers, to improve the understanding of complex emotional expressions. The system uses complex, constantly updated data sets and tries to more accurately capture evolving linguistic nuances and context-specific emotions. In addition, implementing explainable AI methodologies ensures transparency and interpretability in the decision-making process and addresses concerns related to model bias and opacity.
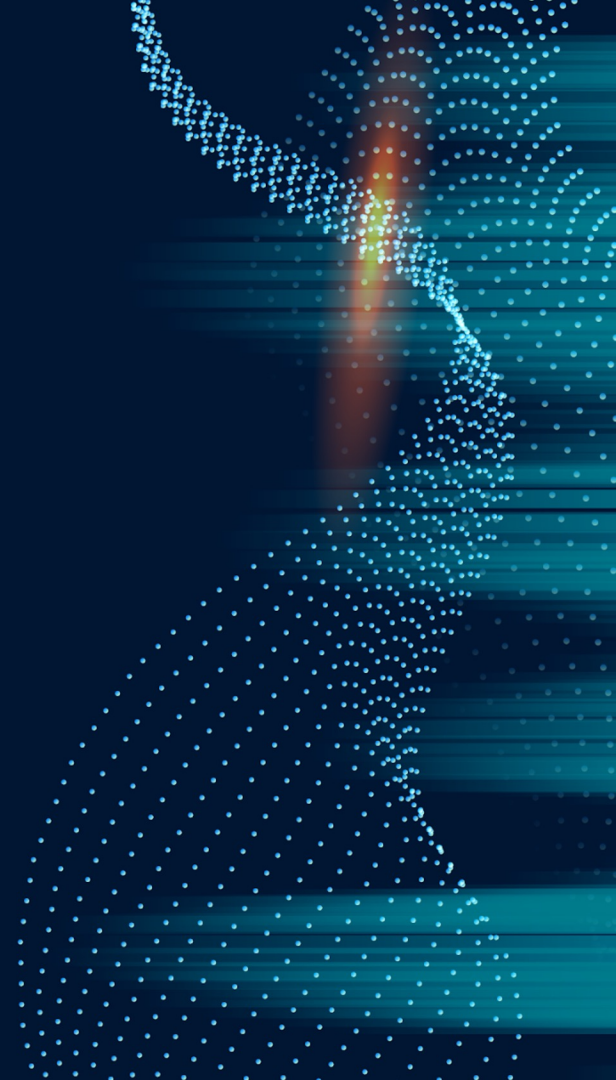
SYSTEM ARCHITECTURE
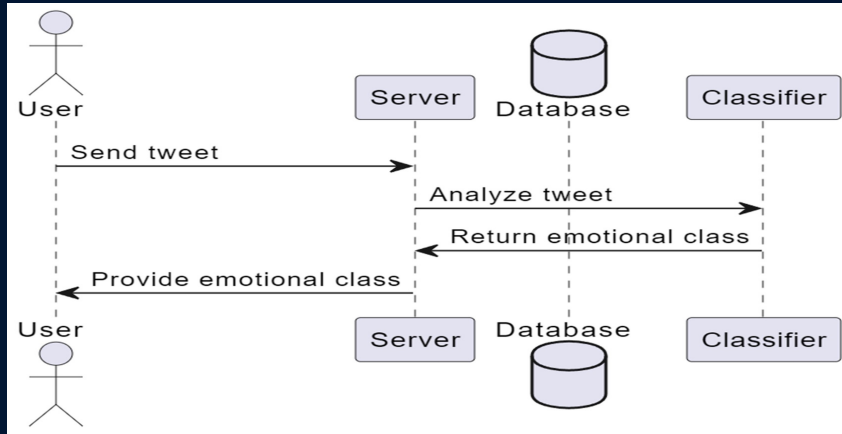
This diagram represents the flow of analyzing textual tweets for emotional recognition. The User sends a tweet to the Server, which then forwards it to the Classifier for analysis. The Classifier determines the emotional class of the tweet and returns it to the Server. Finally, the Server provides the emotional class to the User.

MODULE DESCRIPTION

- **Data Collection Module:** This module involves the systematic gathering of textual tweets from various social media platforms, incorporating efficient web scraping techniques and API integrations to compile a diverse dataset representative of a wide range of emotional expressions and linguistic variations.

- **Preprocessing Module:** In this module, the collected textual data undergoes a series of preprocessing steps, including text normalization, tokenization, and the removal of noise and irrelevant characters, ensuring the standardization and optimization of the textual content for further analysis.

MODULE DESCRIPTION

- **Feature Extraction Module:** This module focuses on extracting relevant features from the preprocessed textual data, utilizing advanced linguistic analysis and sentiment scoring methods to identify key patterns, contextual cues, and emotional indicators essential for the accurate classification of diverse emotional states.

- **Emotion Classification Module:** Leveraging sophisticated deep learning architectures and sentiment analysis techniques, this module categorizes the extracted features into distinct emotional categories, enabling the precise identification and classification of emotions, such as happiness, sadness, anger, and fear, among others, based on the processed textual data.

REFERENCES

- [1] C. Kariya and P. Khodke, "Twitter sentiment analysis," in Proc. Int. Conf. Emerg. Technol. (INCET), Jun. 2020, pp. 212–216.

- [2] A. Alsaeedi and M. Zubair, "A study on sentiment analysis techniques of Twitter data," Int. J. Adv. Comput. Sci. Appl., vol. 10, no. 2, pp. 361–374, 2019

- [3] D. Gamal, M. Alfonse, E.-S. M. El-Horbaty, and A.-B. M. Salem, "Twitter benchmark dataset for arabic sentiment analysis," Int. J. Modern Edu. Comput. Sci., vol. 11, no. 1, pp. 33–38, Jan. 2019.

- [4] Srishti Vashishtha, Seba Susan, "Fuzzy rule based unsupervised sentiment analysis from social media posts", Expert Syst. Appl. 138 (2019) 1–15.

- [5] Peng Wu, Xiaotong Li, Si Shen, Daqing He, "Social media opinion summarization using

REFERENCES

- [6] Muhammad Asif, Atiab Ishtiaq, Haseeb Ahmad, Hanan Aljuaid, Jalal Shah, "Sentiment analysis of extremism in social media from textual information", Telemat. Inform. 48 (2020) 1–20.

- [7] R. Nithya, K. Amudha, A.S. Musthafa, D.K. Sharma, E.H. Ramirez-Asis, P. Velayutham, V. Subramaniyaswamy, S. Sengan, "An optimized fuzzy based ant colony algorithm for 5G-MANET", Comput. Mater. Continua 70 (1) (2021) 1069–1087.

- [8] R. Bharti, A. Khamparia, M. Shabaz, G. Dhiman, S. Pande, P. Singh, "Prediction of heart disease using a combination of machine learning and deep learning", Comput. Intell. Neurosci. (2021).

- [9] Wingyan Chung, Daniel Zeng, "Dissecting emotion and user influence in social media communities: an interaction modeling approach", Inf. Manag. 57 (1) (2020) 1–49

- [10] Kashfia Sailunaz, Reda Alhajj, "Emotion and sentiment analysis from Twitter text", J. Comput. Sci. 36 (2019) 1–18.

THANK YOU