

Hybrid Approaches Based Emotion Detection in Memes Sentiment Analysis

T. Nikil Prakash¹, A. Aloysius²

¹Research Scholar, Department of computer science, St. Joseph's College, Thiruchirappalli, affiliated to Bharathidasan University, Thiruchirappalli.

²Assistant Professor, Department of computer science, St. Joseph's College, Thiruchirappalli, affiliated to Bharathidasan University, Thiruchirappalli.

Abstract

A social network is a decent platform in which users share daily emotions and opinions. As a consequence, they have become an essential information source related to sentiment/opinion. The purpose of Sentiment Analysis (SA) techniques are used to extract sentiments, emotions, and opinions from texts. This data is available by different data sources, such as social media, e-blogs, e-sources, etc. The problem of previous research work has done only text or image analysis using different techniques. Most of the research work results have given low performance. This research work presents a hybrid approach that is needed for the computational processing of Internet memes. Internet Memes provide different behavior of human expressions and activities. The increasing prevalence of Internet memes on social media sites such as Facebook, Instagram, and Twitter also indicates that such multimodal content can no longer be ignored. The present work is analyzing internet memes using lexicon-based and machine learning approaches. In sentiment analysis, every meme is denoted as different polarities (positive or negative memes). The emotions contain six classes that are anger, fear, happiness, sadness, disgust, and neutral. The polarity contains positive (+1), neutral (0), and negative (-1) values. Additionally, the results obtained from the proposed work give better performance of sentiment analysis. The emotion and polarity classes are predicted in both text and image data

Keywords: Hybrid approaches, Sentiment analysis, Memes, Emotions, Convolutional Neural Network. Lexicon-based approaches, Machine Learning,

I. INTRODUCTION

The emergence of social media has opened new possibilities for better understanding the desires of individuals in subjects, brands, or goods. Users of social media constantly upload pictures along with their views and express their thoughts. This trend has enabled the growth of new application areas for machine learning, such as the selection of semantic images from Social Event Analysis, and Emotion Analysis on Visual Contents (i.e. image, audio, videos).

Facial expression recognition is one of the most relevant variables for human expression beings to play the role of interpreting their feelings and motives in interaction.

Encompassing a vast range of data and carrying emotional meaning, facial expressions become one of the important data networks in interactive communication [1]. Visual Sentiment Analysis attempts to infer the sentiment evoked by pictures. In this area, early approaches concentrated only on visual features (ignoring the text associated with the images) or used text to describe the ground truth of a feeling. More recent methods take advantage of a mixture of visual and text characteristics in various ways. Most of them, consider SentiWordNet lexicons as external knowledge to extract useful semantic information from textual data [2].

The increasing ubiquity of Internet memes on social media sites such as Facebook, Instagram, and Twitter has become a subject of tremendous interest over the past few years. Memes, one of the most typical English phrases. Memes are also extracted from our previous social and cultural interactions, such as TV shows or a famous character of a cartoon. As a meme may be both a textual representation and a visual representation, a meme is uniquely multimodal. It may also be used in a multimodal sense where, along with textual comments and interactions, the meme (whether with text or not) is used to provide focus, additional meaning, or an effective response [3-4].

II. RELATED WORK

Dibyedu seal et al. [5] presented the phrase-level identification of emotions in words and phrasal verbs. The Emotion Database contains a large number of words that directly affect emotions. It is difficult to draw up a comprehensive list of words, many words are indirect emotions. The Emotion Detection Method described the Emotion Database process in sentence-level words and phrases.

Ekaterina Ivanova et al. [6] proposed the enhancement of the method of recognition of emotional expressions in the human face. This method is based on a sentimental recognition system based on deep neural networks. Emotional recognition techniques were used to evaluate public and private text samples. Each group of emotions provides the highest accuracy of image recognition. This model has been evaluated from neural networks.

Xinzhi Wang et al. [7] suggested the relation between expression of emotion in behavioral uncertainty and analysis. They mainly focused on human facial emotions. For trading emotional trends, emotional correlation is important.

Three types of features and two deep neural networks were created for this model. This model was evaluated by subjective and objective methods and obtained through long web text notes.

Jessy Jane Christy et al. [8] proposed emotion detection used in psychotherapy treatments. It identifies and recommends people with severe or mild mental disorders based on symptoms in the form of text or speech. To classify the feelings, they described different approaches. These approaches are namely Convolutional Neural Networks, Support Vector Machine, Gradient Descent Algorithm, Naïve Bayes, Long Short Term Memory, etc. the emotion detection method can do in two steps that are feature extraction and classification. Feature extraction involves extracting prosodic and phonetic features from speech. The classification is to classify mathematical neural networks by speech.

Nodine Brawn et al. [9] suggested a miscommunication or misunderstanding of the emotions conveyed through the text. If the emotion is expressed through the text. It doesn't detect the matched ones. The automatic method for detecting emotions is addressed in the sentiment analysis system. Approaches to analyzing the effects of the text were considered in this method.

Jon Hofmann et al. [10] proposed automatic emotion detection classification in text. These approaches need to learn the proportion of events as latent variables. The cognitive appraisal method represents identifying the situation and the emotion classification model classified the text-based method.

Carlo Lipizzi et al. [11] described the text analysis method of subjectivity and content as greater reliance and detected emotions in the text being conveyed. This method has classified subjectivity and text vectorization using Word2Vec approaches. This method classified three categories, namely the creation of a point of view, the standard measure of basic emotions, and the document analysis. This method is evaluated in the case of text in political sources.

John Kalung et al. [12] proposed the concept of user emotion vectors and film emotion vectors. This method describes various approaches to build emotions that are content-based and collaborative filtering methods and an item-based collaborative filtering method, a user-based collaborative method, a content-based emotion-conscious recommendation, and a content-based multichannel affection recommendation

III. PROPOSED WORK

Multimedia involves multiple types of text, audio, images, video, etc. Twitter has steadily converted from a text-based micro-blogging (very micro) platform to a visual one. The proposed method takes into consideration the text, and images separately and the mixture of these to examine emotions in tweets. The incoming tweet is analyzed first of all for its methodology form i.e. whether it is an image, text, or multimodal text (image + text = typographic or info-graphic). The proposed work categories into three-level. These are explained below section. The first section is applying preprocessing techniques second section is applying feature extraction methods for processed data and finally, the emotions and text are classified in this section.

The Senti_Emotion_Hybrid method is proposed in this research work. The Senti_Emotion_Hybrid method is analyzed both textual and face data. The Senti_Emotion_Hybrid method combine lexicon-based approaches and Machine Learning approaches. Fig 1. Shows the framework for the proposed work model.

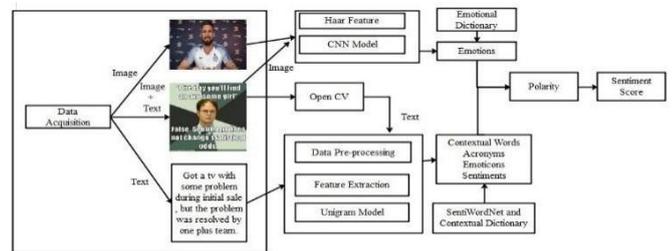


Fig. 1. Architecture for proposed work sentiment analysis

Level 1: Preprocessing

a. Data Collection

There are many open-source datasets are available for memes, like Facebook, Twitter, WhatsApp, etc. Also, there are various open-source datasets offered for memes. The dataset is collected from the Kaggle website. Memes dataset is collected from the proposed work. Memes contain three different types in sentiment analysis, namely image, text, and image-text.

b. Data Preprocessing

Data preprocessing is the process of cleaning and preparing the data for classification. Online data contains a lot of noise and uninformative parts such as Html tags, advertisements, and scripts. The proposed data preprocessing techniques remove the noise data, documents, videos, audios, tags, etc. The text data is analyzed by lexicon-based approaches [24], face data is analyzed by machine learning approaches. Finally image-text data, the text is extracted using the OCR method after extracting the text is analyzed by the lexicon-based approaches then the image is applied by machine learning approaches.

Level 2: Feature Extraction

The feature extraction method is used to extracting the features of text and image data. The text features are extracted using the lexicon-based feature method and the image features are extracting using machine learning feature methods. The lexicon features are acronyms, emoticons, and contextual words are applying and extracting the text features. The Haar-Like feature and Convolutional Neural Network model features are extracting the face features in image data.

A. Lexicon based approaches: Textual feature extraction

Textual feature methods contain acronyms, emoticons, and contextual words. These features are extracted by lexicon dictionary that is SentiWordNet and Contextual Dictionary. The Senti_Con_Acro algorithm [22] is extracting these textual features. Text analysis also contains four levels of analysis. These are performed on textual comments that are namely word-level analysis, phrase-level analysis, comment analysis, and discussion analysis.

B. Machine Learning Approaches: Facial Expression Analysis

Facial expression analysis is the process of analyzing whether the face is determining positive or negative categories and or detecting emotions (happy, sad, anger, etc.). In this proposed work the Haar-Like feature method is extracting the faces in images and cropping the faces in image data. Then applying Convolutional Neural Network model layers (dense, dropout, conv2d, etc.). This model is extracting the facial features in the faces. The Senti_Emotion [23] method is used for extracting facial features in face data.

C. Hybrid Approaches: Image and Text Analysis

It is identifying the meaning of the meme using both the images and text. The main focus of meme analysis is extracting the text and facial features. The Optical Character Recognition (OCR) method is used for extracting the text in images. Fig 1 shows extracting the text in the image.

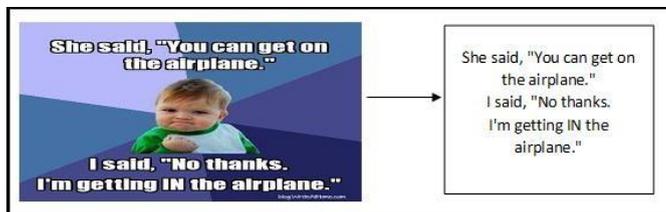


Fig 2. Example of extracting the text in the image

The meme analysis contains two different categories that are namely 1) meme text descriptions which are described in the textual phrase, 2) meme categories contains descriptive tags such as happy, sad, upset, etc.

D. Optical Character Recognition (OCR)

The Proposed open-source optical character recognition engine is one of the most effective and most powerful OCR methods

in machine learning approaches. OCR is used to search and detect text on images. OCR is used in pixels, shapes, words, and phrases to classify representations. After the extracted text is passed to the data pre-processing techniques.

Level 3: Classification

A. Emotion Classification

After the extracted facial features are represented as six emotions such as happy, sad, angry, fear, disgust, and neutral. These are classified as each facial expression.

B. Polarity Detection

The text and emotions are classified by the polarity detection method. The polarity detection method is to find the polarity value of text sentiment and facial emotions. The happy and surprise emotions are denoted as positive categories and the sad, anger, fear, disgust emotions are denoted as negative categories and the other emotions are denoted as neutral categories. All the positive, negative, and neutral values are evaluated by the confusion matrix method that is precision, recall, f-measure, and accuracy.

IV. RESULTS

The Memes dataset is collected from this proposed work. The memes dataset 44800 memes (both text and images) are classified and 35898 memes are identified. Textual data is analyzed by lexicon-based approaches using the Senti_Con_Acro method. Facial expression emotions are classified as Senti_Emotion approaches. The Emotion classes are affecting anger, happiness, sadness, fear, disgust, and neutrality. The model computes positive, negative, and neutral emotions and sentiments. The confusion matrix method evaluates the proposed Senit_Emotion_Hybrid model. The Senit_Emotion_Hybrid model obtains better performance. In "Softmax" function is used to reduce the losses in the memes dataset.

Table 1. The experimental results for the confusion matrix

	Precision (%)	Recall (%)	F-Measure (%)	Accuracy (%)
Proposed Result	79.81	84.88	82.27	86.53

Table 2. The comparison results for existing and proposed Work

Authors	Accuracy (%)
Seunghyun Yoon	65.5
Medha Khurana	75
Hazarika	77.6
Mohammad Faridul	82.26
Proposed Work	86.53

A. Confusion Matrix Method

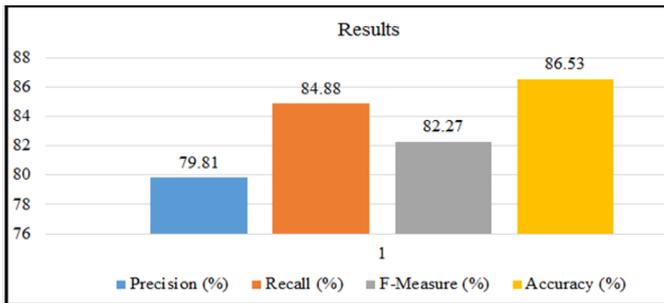


Figure 3. The experimental results for the confusion matrix.

V. CONCLUSION

The study of feelings in Memes is comparatively more complicated because the meme is more conceptual and subjective. Meme analysis has two types of information that are text and visual (image, video, gif, or flash files). This study describes the method of emotional detection by identifying sentiment and emotions from text and facial expressions using the memes dataset. The Senti_Emotion_Hybrid model improves the performance of existing methods. Also, classified lexicon-based and machine learning techniques. In the future, we will analyze different levels of texts and emoji's in memes. Additionally, we will try to analyze gif, cartoon, and animation memes with different techniques and features.

REFERENCES

- [1] Joanna Hong, Hong Joo Lee, Yelin Kim, and Yong Man Ro "Face Tells Detailed Expression: Generating Comprehensive Facial Expression Sentence through Facial Action Units" 26th International Conference, MMM 2020, DOI: https://doi.org/10.1007/978-3-030-37734-2_9, 2020, 100–111.
- [2] Alessandro Ortis, Giovanni Maria Farinella, Giovanni Torrisi, Sebastiano Battiato "Exploiting objective text description of images for visual sentiment analysis" Springer, Multimedia Tools and Applications, DOI: <https://doi.org/10.1007/s11042-019-08312-7>, 2020, 1-24.
- [3] William Theisen, Joel Brogan, Pamela Bilo Thomas, Daniel Moreira, Pascal Phoa, Tim Weninger, Walter Scheirer "Automatic Discovery of Meme Genres with Diverse Appearances" Association for the Advancement of Artificial Intelligence, 2020, 1-13.
- [4] Jesus Perez-Martin, Benjamin Bustos, Magdalena Saldana "Semantic Search of Memes on Twitter" Computational Methods Interest Group of the 70th International Communication Association Conference, 2020, 1-14.
- [5] Dibyendu Seal, Uttam K. Roy and Rohini Basak "sentence-level Emotion Detection from Text Based on Semantic Rules" Information and Communication Technology for Sustainable Development, Advances in Intelligent Systems and Computing 933, DOI: https://doi.org/10.1007/978-981-13-7166-0_42, 2020, 423-430.
- [6] Ekaterina Ivanova and Georgii Borzunova, "Optimization of machine learning algorithm of emotion recognition in terms of human facial expressions" Postproceedings of the 10th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA 2019 (Tenth Annual Meeting of the BICA Society), 169: 2019. 244–248.
- [7] Xinzhi Wang, Luyao Kou, Vijayan Sugumaran, Xiangfeng Luo, and Hui Zhang "Emotion Correlation Mining Through Deep Learning Models on Natural Language Text" IEEE Transactions on Cybernetics, 2020, 2168-2267.
- [8] J Jessy Jane Christy, S Iwin Thanakumar Joseph, S Sudhakar, V.M Arul Xavier, "Emotion Detection in Text and Acoustic Based Applications" Journal of Green Engineering, 10_6: 2020. 3006–3020.
- [9] Nadine Braun, Martijn Goudbeek and Emiel Kraherer "Emotional Words – The Relationship of Self- and Other-Annotation of Affect in Written Text" Cognitive Science, 2020, 1-7.
- [10] Jan Hofmann, Enrica Troiano, Kai Sassenberg, and Roman Klinger "Appraisal Theories for Emotion Classification in Text" Computation and Language, DOI: arXiv:2003.14155v4 [cs.CL], 2020, 1-14.
- [11] Carlo Lipizzi, Dario Borrelli, Fernanda de Oliveira Capela "A computational model implementing subjectivity with the "Room Theory" – The case of detecting Emotion from Text" Computation and Language, DOI: arXiv: 2005.06059, 2020, 1-15.
- [12] John Kalung Leung, Igor Griva and William G. Kennedy "TEXT-BASED EMOTION AWARE RECOMMENDER" DOI: 10.5121/csit.2020.101009, 2020, 101-114.
- [13] Yazhou Zhang, Dawei Song, Xiang Li, Peng Zhang, Panpan Wang, Lu Rong, Guangliang Yu, Bo Wang "A Quantum-Like multimodal network framework for modeling interaction dynamics in multiparty conversational sentiment analysis" Elsevier, Information Fusion, 62, 2020, 14–31.
- [14] Seunghyun Yoon, Subhadeep Dey, Hwanhee Lee and Kyomin Jung "Attentive Modality Hopping Mechanism for Speech Emotion Recognition" IEEE, ICASSP, DOI: 978-1-5090-6631-5, 2020, 3362- 3366.
- [15] Mohammad Faridul, Haque Siddiqui and Ahmad Y. Javaid "A Multimodal Facial Emotion Recognition Framework through the Fusion of Speech with Visible and Infrared Images" Multimodal Technol. Interact. DOI: 10.3390/mti4030046, 4(46); 2020 1-21.
- [16] Hazarika, D. Poria, S. Zadeh, A. Cambria, E. Morency, L.P. Zimmermann, R. "Conversational Memory Network for Emotion Recognition in Dyadic Dialogue Videos" Proceedings of the 2018 Conference of the North

American Chapter of the Association for Computational Linguistics: Human Language Technologies, USA, 1: 2018, 2122–2132.

- [17] Liang, P.P, Zadeh, A. Morency, L.P. “Multimodal local-global ranking fusion for emotion recognition. In Proceedings of 2018 on International Conference on Multimodal Interaction, ACM: 2018, 472–476.
- [18] Livingstone, S.R. Russo, and F.A. “The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English. PLoS ONE, 2018, 1-13.
- [19] Lee, C.W, Song, K.Y.; Jeong, J, Choi, W.Y. “Convolutional Attention Networks for Multimodal Emotion Recognition from Speech and Text Data” arXiv, arXiv:1805.06606. 2018.
- [20] Majumder, N, Hazarika, D, Gelbukh, A, Cambria, E, Poria, S. “Multimodal sentiment analysis using hierarchical fusion with context modeling” Knowl- Based Syst, 161: 2018, 124–133.
- [21] Seunghyun Yoon, Seokhyun Byun, Kyomin Jung “Multimodal Speech Emotion Recognition Using Audio and Text.” IEEE Spoken Language Technology Workshop (SLT), DOI: 10.1109/SLT.2018.8639583, 2019, 112-118.
- [22] T. Nikil Prakash and A. Aloysius, “Lexicon Based Sentiment Analysis (LBSA) to Improve the Accuracy of Acronyms, Emoticons, and Contextual Words”, Society of Statistics, Computer and Applications, (Accepted-Forthcoming), 2021.
- [23] T. Nikil Prakash and A. Aloysius, “Facial Expression based Emotion Detection in Sentiment Analysis”, Punjab University Journal of Mathematics, (Communication) 2021.
- [24] T. Nikil Prakash, A. Aloysius “Applications, Approaches, and Challenges in Sentiment Analysis (AACSA)”, International Research Journal of Modernization in Engineering Technology and Science, ISSN: 2582-5208, Vol. 2, Issue 7, pp. 910-915, 2020.