https://doi.org/10.46335/IJIES.2022.7.8.16 e-ISSN: 2456-3463
Vol. 7, No. 8, 2022, PP. 89-96

International Journal of Innovations in Engineering and Science, www.ijies.net

Sentiment Identification from Image-Based Memes Using Machine Learning

Amit Pimpalkar[1], Ayush Chaudhari [2], Atul Lilhare [3], Nimish Dighorikar [4], Sagar Dakhole [5], Shreyash Asawa [6]

[1]Assistant Professor, [2,3,4,5,6] Scholars
Department of Computer Science and Engineering, G H Raisoni Institute of Engineering and Technology, Nagpur

amit.pimpalkar@raisoni.net

***Received on**: 11June ,2022* ***Revised on**: 31 July ,2022,* ***Published on**: 03 August,2022*

***Abstract-** People increasingly share a lot of content on social media intending to discover memes that express their feelings. Image and textual Sentiment Analysis (SA) is gaining pace as a result of the rising tendency of expressing sentiments through images and written descriptions. Photographs and videos are increasingly being used by social media users to express themselves and share their experiences. In this study, we attempt to anticipate the meaning of memes, such as whether they are good, negative, or neutral. This will help us understand how individuals feel about a certain topic. To do this, we developed unimodal and multimodal techniques for extracting visual and linguistic information from memes. In our dataset, we employed the Bidirectional Encoder Representations from Transformers (BERT) pre-train model for text categorization.*
.

Keywords- Sentiment-Analysis, Memes, OCR, BERT, Image Captioning,VGG-16

I – INTRODUCTION

Memes on social media sites such as Facebook, Instagram, and Twitter have become a subject of great interest in recent years as the Internet has grown in popularity. Memes have become widespread internet content over the last decade, often via social media platforms and especially for entertaining causes. Memes can also be derived from earlier social and cultural exchanges, such as television episodes or a well-known cartoon character. A meme is uniquely multimodal in that it can be both a textual and a visual representation.

Memes are no longer simply for fun; they also draw attention to cultural and social concerns. In other areas of human effort, intellect is the driving force behind going viral.With the growth of social media, new channels for better understanding people's preferences in problems, businesses, and commodities have opened up. Social media users commonly share memes/pictures along with their ideas and thoughts.This trend has aided in the analysis of their ideas, thoughts on a specific topic, and opinions on a specific topic through Memes/Pictures shared on social media utilizing Machine Learning (ML), i.e., SA, emotion analysis, or opinion mining. The process of defining and categorizing opinions in a given piece of text as positive, negative, or neutral is known as SA, also known as opinion mining.

One of the most important variables for human expression beings to interpret their sentiments and motives in contact is facial expression recognition Facial expressions have become one of the most important data networks in interactive communication, encompassing a wide range of data and carrying emotional value. Visual SA tries to deduce the emotion elicited by images. Early methods in this field either focused solely on visual aspects (thus neglecting the text linked with the images) or used words to express the underlying truth of a sensation. More contemporary methods employ a combination of visual and textual qualities in a variety of ways. It can also be used in a multimodal context, where the meme (whether with text or without) is used to provide focus, additional meaning, or an effective response in addition to textual comments and interactions. Analysis of sentiment is a topic in NLP that has received a lot of attention. Many research papers exist in which a labeled dataset, primarily textual, is used for SA of social media posts, discussions, comments, and topic opinions to determine the deeper context of people's perception and behavior online.

With the advancement of technology, emotion analysis is becoming a more popular tool for businesses and their impact on society.SA is a technique for analyzing customer feedback and reactions to find negative remarks and reasons why customers are dissatisfied with your product or service. SA is extremely valuable in social media monitoring because it provides a wide picture of public opinion on a given issue. SA can be applied in a variety of ways. The ability to derive insights from social data is being embraced by companies all around the world.It's easier to plot and plan for the future when you can easily see the sentiment behind everything from forum postings to news items. It can also assist you with customer service and market research. You may not only see what people think about your products or services, but you can also see what they think about those of your competitors. SA can provide you with a brief glimpse of your customers' overall experience, but it can also be far more detailed. Emotional customer responses such as “positive,”“negative,”“neutral,”“negative,”“uptight,”“dis gust,”“frustration,” and others are referred to as “emotion-based marketing”. Understanding the psychology of customer reactions can also aid product and brand recall. Businesses can only make cultivated guesses, and most of the time, those assumptions can be inaccurate.

In the commercial world, SA may help you scan through hundreds of product evaluations and extract relevant and meaningful information to assess whether or not your customers are satisfied. Techniques for SA make it easier to see the industry through the eyes of customers. Industries can rapidly figure out what's bothering their customers and why they're unhappy with their goods or service. The top-of-the-page thread of unfavorable comments gives you plenty of opportunities to answer and listen to your customers. Because human language is so complicated, your social listening must be able to decipher it to discover emotional phrases. NLP can recognize slang and pop culture phrases, as well as memes and images making image analysis just as vital as the others. In our increasingly graphic online environment, this is even more true. Images are frequently used to replace text and understanding people's psychology can aid content creators and social media influencers in persuading them to change their thoughts about a topic they promote via memes. As this trend continues to grow, the importance of assessing sentiment in photos has grown in lockstep. We are aiming to investigate the relationship between meme content and social network features. Our main goal is to decipher memes from text and images.

The remainder of the paper is arranged as follows: Sect. 2 presents a comprehensive literature review on DL approaches. Section 3 delves into the proposed work's methodology and materials, including dataset collection, preprocessing, OCR, embedding model, and deep neural networks. Section 4 demonstrates the experimental analysis, findings, and comparison with the state of the art, as well as an explanation of the acquired results and their evaluation. Section 5 concludes the paper with additional possibilities for improvement.

II -LITERATURE REVIEW

SA in the multimedia domain is an enormously researched subject matter. Social media platforms are flooded with textual, visible, and multimodal information. That being said, very less attention has been given to the multimodal and visible sentiment evaluation compared to textual content primarily based on sentimental evaluation[14].They introduced Memesem, a multimodal containing deep neural framework network for meme and sentiments analysis using transfer learning. Authors used VGG-16 which they pre-trained on BERT Language modal and Image net dataset to learn textual and visual features of memes and combine them to give results. The model not only performs better than the uni-model but also thrives on baseline multimodal based on text and visuals. The [22] also consider both bimodal (text and image) and unimodal (text-only) strategies, ranging from the Naive Bayes classifier to Transformer-based algorithms. The findings reveal that a text-only approach, using a simple Feed Forward Neural Network with Word2vec embeddings as input, outperforms all others. In [8] the topic of recognizing hate memes was addressed using two alternative approaches. The first method employs SA based on image captioning and meme text. And alternatively, was to mix features from several modalities. These methods combine glove, encoder-decoder, and Optical Character Recognition (OCR) algorithms with Adamax optimizer deep learning

https://doi.org/10.46335/IJIES.2022.7.8.16 e-ISSN: 2456-3463
Vol. 7, No. 8, 2022, PP. 89-96

algorithms, and the findings for both models were found to be promising.

Researchers[1] discovered that sequence models such as Long Short Term Memory (LSTM)[7] and its variants performed poorly in predicting sentiments compared to transformers and they employed the pre-train model BERT for text categorization and it outperformed other Distil BERT and XL Net Transformer-based models, except for motivational sub-tasks where Distil BERT performed well. In research [12]authors compared the models namely VGG-16, ResNet-50, InceptionV3, and MobileNetV2 for sentiment categorization, and in their study, they found that the better performance for the SA was given by VGG-16.Authors proposed that [9], there was no correlation between the photos and the text. When employed with the predictive models that they used in their study, multimodal solutions did not demonstrate significant gains, but employing only one modality (text or image) tends to produce superior outcomes the same. Researchers[18] proposed that the multimodal and unimodal both are not reliable and have debatable due to their accuracy. The text classifier's accuracy is comparable to that of the multimodal classifier. The visual classifier, on the other hand, has a lower chance of detecting and maintaining offending memes on its .A study [11] presented a method for detecting emotions from facial expressions and text by proposing meme datasets to determine sentiment and emotions. Existing approaches perform better with the Senti Emotion Hybrid model, ML, and categorized lexicon-based approaches. Authors compared relevant papers on Image Analysis and their sentiments [2].In work [10], they proposed the uses of visual and textual modern models.

In [21],they investigated and expanded on numerous training models for simple text classification algorithms and discovered that the Convolution Neural Net (CNN)-LSTM architecture provided the best accuracy. Rather than the traditional Subjective Text provided by users, [3]their model extracts and uses object text of Image spontaneously from the visual information. The model used in this method was multimodal embedded space based on textual features. The [20] offer a clever model based on a transformer architecture that attempts to achieve state-of-the-art by utilizing attention as its primary component, Troll and non-troll Tamil meme images were included in the collection, along with their captions as text. LSTM was used in the proposed model for understanding complex patterns in textual data[19].To improve the performance of the LSTM, weight parameters were optimized by the adaptive Particle Swarm Optimization algorithm and they found that the algorithm APSO-LSTM has outperformed LSTM, ANN, and SVM [4-5]. To improve the F1 score authors have created their model by concatenating the solution. Research of [15] suggested a method for detecting hateful memes that won third place in the Hateful Meme Challenge, which uses a pre-trained Visual BERT fine-tuned on an enlarged train dataset.

Authors of [13] compared multiple bi-fusion techniques for meme emotion analysis, and the RoBERTa+ResNet models were found to be superior. A method had been proposed to detect offensive memes by [16]. In this model, they classified memes into two categories i.e. First offensive and second non-offensive after extracting text from the image. The model has simple architecture containing multilayer dense network structure involving NLP with RNN and LSTM, as well as GloVe[7] and rapid text word embeddings.

III -METHODOLOGY

We are inspired by the work of both text and visual models (text and images) by the previous work. Since we have memes as our input, we are taking both unimodal and multimodal approaches. We generate a caption of an image meme and then we extracted text from a meme, the extracted text is preprocessed, and then the clean text and image caption are generated, and then the concatenated text is sent to the BERT model for the text classification to generate the score for the meme sentiment. We use the same methods for unimodal as we do for multimodal, only we don't generate image captions and only the extracted text from memes is preprocessed before being given to the BERT model for text classification to generate the meme sentiment score.

Deep Learning Formulation of the Business Problem

Text Detection: Text Detection is a technique in which a picture is fed into the model and the textual region is recognized by drawing a bounding box around it.

Text Extraction: Following text detection, text recognition occurs, in which the discovered textual portions are further processed to recognize and extract what the text is [17].

Image Segmentation: Image segmentation is a way of breaking down a digital image into several subgroups called Image segments to reduce the complexity of the image and make future processing or analysis of the image easier. At each level, various deep learning models were deployed to execute the aforementioned tasks.

https://doi.org/10.46335/IJIES.2022.7.8.16 e-ISSN: 2456-3463
Vol. 7, No. 8, 2022, PP. 89-96

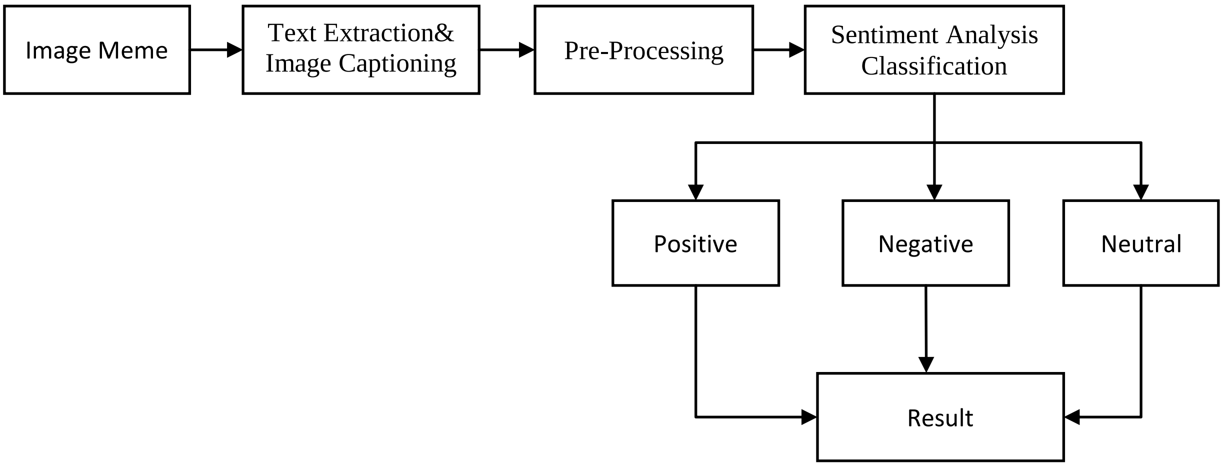

Fig 1- Stages of Sentiment Analysis for memes

A SA can be discriminated against on the following types.

Positive: Positive words or sentences contain these feelings. Some positive literature conveys feelings of joy, friendliness, and enthusiasm, among other things. For instance, he is a genius, and this view is breathtaking, and so on.

Negative: Negative attitudes can be found in textual or visual content in the form of negative phrases or sentences. Some bad emotions are hated, grief, violence, and so on. negative comments have the option to be reported or blocked. For example, this view is unacceptable, he is my adversary, and so on.

Neutral: Some statements may have no emotion, which is referred to as neutral sentiment. For instance, there is a novel on the couch, I am eating, and so on.

Dataset:

The Facebook Hateful Meme Dataset [23] is an open-sourced dataset containing offensive memes on Kaggle. Memotion Dataset [24] is a Kaggle-hosted open-source dataset containing memes with sentiments positive, negative, and neutral, which includes labels that have been annotated by humans. Following an observational review of the dataset, we claim that the bulk of the samples is incorrectly categorized. There had been no freely released dataset for detecting meme sentiment; there were some, but they were inaccurate or didn't include memes with all three types of sentiment: positive, negative, and neutral. As a result, we built our datasets in which memes were retrieved from Reddit, Instagram, and Facebook. So, the memes were filtered by the following guidelines: The meme must include a clear background image as well as textual material incorporated in it. For this study, only memes with English language text content are considered.

Table 1- Memes Dataset Statistics

Dataset	Total	Positive	Negative
Reddit + Instagram meme	743	438	305
Memotion + Facebook Hatefull meme	6873	4089	2784

Typical steps for loading custom datasets for Deep Learning Models

- **Open the image file**. The format of the file can be JPEG, PNG, etc.
- **Resize the image to match the input size for the Input layer of the Deep Learning model**.
- **Convert the image pixels to float datatype.**
- **Normalize the image** to have pixel values scaled down between 0 and 1 from 0 to 255.
- **Image data for Deep Learning models should be either a NumPy array or a tensor object.**

OCR:

One of the most researched fields in AI and DL is optical character recognition. Many academics have completed many model designs, but none of them can be generalized because it all relies on the dataset they are using. This gives us a wide grasp of how we might solve comparable challenges. For our experiential purpose to process memes separately, we need to extract the text data from them. For which Tesseract-OCR is used. It's a free and open-source OCR engine for extracting text from images based on LSTM networks. Tesseract-OCR extracts the meme dataset's embedded texts. These texts

https://doi.org/10.46335/IJIES.2022.7.8.16 e-ISSN: 2456-3463
Vol. 7, No. 8, 2022, PP. 89-96

were saved in CSV files for future use. It works well on x86/Linux with an official Language Model and data available for 100+ languages and 35+ scripts.

Image Captioning:

Image captioning is a growing study area. It is quite difficult to extract context from an image just by looking at it. To address such issues, DL encoder and decoder designs are used. The task of Image Captioning can be broken into two parts:

1. Model-based on images: To get the feature out of the picture.
2. Model-based on Language: To convert our image-based model's derived features into a natural sentence.

For image classification a well-known pre-trained CNN model is VGG-16 and we employ LSTM for the model based on language.

We have combined the architectures of CNN and LSTM for Image Classification.

The model (VGG-16+LSTM) is trained on the [25] Flickr dataset. With 20 Epochs, Batch-size of 32, and Optimizer adam.

Data Preprocessing:

The text produced may contain a lot of noise, such as the name of the meme page, as well as incorrect text extraction and thus they must be corrected by us. We've discovered contracted terms such as we've, isn't, and so on and replaced in this case, we have, is not, etc. To simplify all of the terms, we used tokenization and lemmatization and this was accomplished with the help of the NLTK and genism Python packages. Tokenization divides the text into a list of words, lowercase all letters, and removes punctuation. All verbs appear in the first person, present tense form after being lemmatized. Words with repeating letters, such as much, no, were transformed to much no., '#'is Removed from hash tags, and Stop words with more than three characters are not removed from text to provide sufficient contextual information. Stop words and words with less than three characters were eliminated [4-7].

Text Classification:

We now have the cleaned text, which is ready for training, after the data has been pre-processed. There are two methods for training the data: training from scratch and using the pre-train model. So, according to [8], they compared different pre-train models and found that BERT performed admirably in most meme categories. As a result, we are also using the BERT retrained model. Because BERT works with context, each word is represented as a vector with its tokenizer ID, and also a segment ID to differentiate among sentences as well as an input mask to distinguish among both tokens and padded sequence. Every post's pre-processed data is then passed into a BERT, which extracts textual context. To make final predictions, textual features are passed further into as input in32-dimensional convolution layers that utilizes a SoftMax function combined with the BERT.

VGG-16

VGG-16 is a 16-layer deep CNN architecture andthe most distinguishing feature of VGG-16 is that instead of a huge number of hyper-parameters, they concentrated on having convolution layers of 3x3 filter with stride 1 and always utilized the same padding and maxpool layer of 2x2 filter with stride 2. This configuration of convolution and max pool layers is maintained throughout the design. Finally, it features two completely connected layers and a soft ax for output. Because we are not considering the fully-connected output layers of the model used to generate predictions, we loaded the retrained VGG-16 model with include top = False, allowing a new output layer to be added and trained.

BERT Pre-trained Model :

The BERT algorithm is a natural language processing-related deep learning technique. BERT is intended to assist computers in understanding the meaning of ambiguous words in the text by establishing context from surrounding content.Instead of creating a model from scratch, we may leverage cutting-edge models like BERT to do classification jobs.BERT typically accepts three inputs: tokens, masked tokens, and segmentation tokens.And they are submitted to BERT, and we will eventually acquire a feature vector from which we will design our classifier.

Table 2: Training Function Parameters

Training Parameter	Values
OutputFunction	Sigmoid
BatchSize	32
Shuffle	every-epoch
InitialLearning Rate	1e-4
Max Epochs Number	20
Dropout	0.4
ActivationFunction	Relu, Softmax
OptimizerFunction	Adam
Convolution Layer Filter	3x3, 2x2
Stride	1, 2

https://doi.org/10.46335/IJIES.2022.7.8.16 e-ISSN: 2456-3463
Vol. 7, No. 8, 2022, PP. 89-96

Table 2 lists all of the hyper parameter values utilized in the experiment. The model's hyper parameters are entirely determined by the experiments. We run multiple iterations with all feasible and different hyper parameter setups. After testing the model's performance on multiple hyper parameter settings, we chose the one that produced the best results on our dataset.

IV -EXPERIMENTAL RESULT AND DISCUSSION

In Table 3, the proficient significance of the two models has been particularized. We believed that the image captioning model was the most important factor in achieving a higher F1score,as there is no image captioning dataset for memes. A dedicated meme captioning dataset can greatly improve multimodal accuracy. Still, we also believed that the difference could be greater if the model was trained more thoroughly on a larger dataset. Using the F1 score we discovered that multimodal performed slightly better around 5%than the unimodal on the Reddit + Instagram meme dataset, whereas Memotion + Facebook Hateful meme dataset resulted in no big improvement. The results for the same have been revealed in Figure 2.When compared to human accuracy, we earned a respectable outcome on this topic.

Table 3: Performance Evaluation of proposed model On Dataset With Unimodal And Multimodal

Model	Unimodal			Multimodal		
Dataset	Precision	Recall	F1	Precision	Recall	F1
Reddit + Instagram meme	59.51	68.96	63.89	64.29	72.14	67.99
Memotion + Facebook Hatefull meme	64.28	77.06	70.04	64.41	75.73	69.61

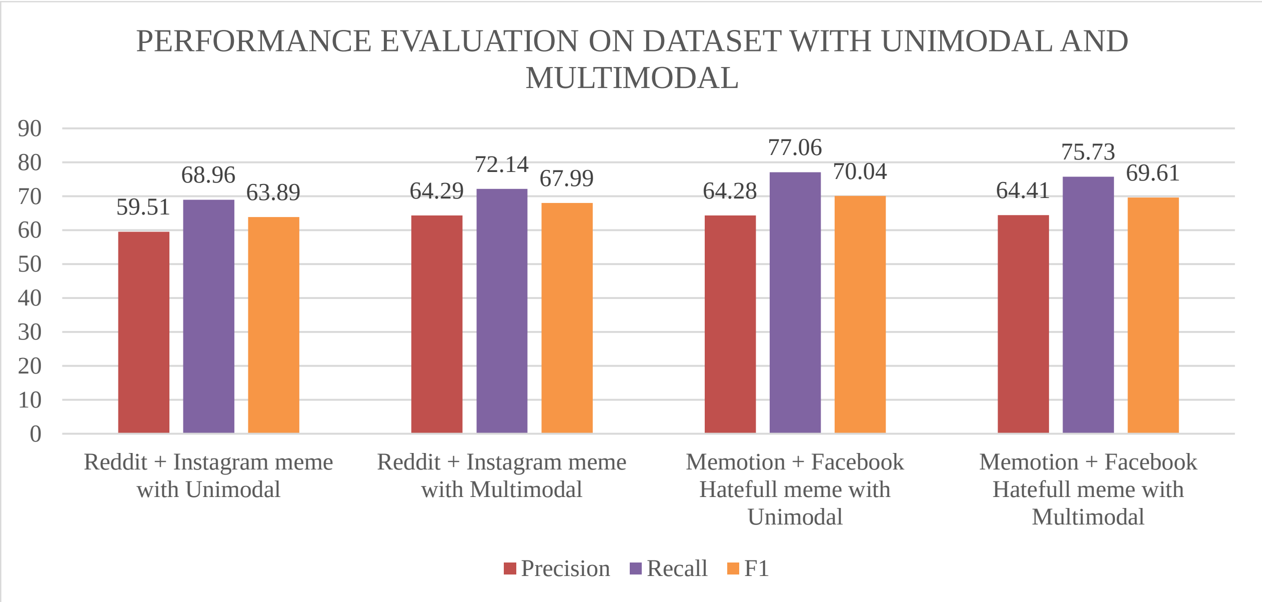

Fig 2- Performance Evaluation On Dataset With Unimodal And Multimodal

V- CONCLUSION AND FUTURE WORKS

In this research, we find out whether the sentiment of the meme is either positive or negative. Coming to its benefits it is useful in social media monitoring and it gives public opinion behind certain topics.We trust that sentiment classification on substantial quantities of internet user-generated material is beneficial since it can provide more reliable signals and information for a variety of data analytics tasks using digital platforms for prediction. As no single model is capable of analyzing all kinds of genres in a meme, as there isa variety of negative memes based on racism, religion, politics, and terrorism and a positive text score does not necessarily imply a positive meme. We believe that this model is still a long way from evaluating humans. The negative meme is not always offensive, if it's used to remove offensive meme content, an invigilator should be present to ensure or double-check. For social media SA, we intend to improve the performance of the models multimodal that use both textual and visual content.

We have meme images, text, and corresponding labels but the sentiment message can be embedded in different layers of image abstraction, visual SA. In the future, we can combine SA across different aspects of humor, sarcasm, and motivation also Hyper parameter adjustment can be used on the current model to increase accuracy. Further , generating a dataset of meme templates and training our image captioning model on it can greatly increase the algorithm's understanding of memes based on those templates. We also intend to use a facial recognition model in future because memes did contain faces, allowing for more accurate multimodal work. We can investigate in-depth with different models to increase the model's performance. We also hope that the results of our SA will inspire more research into meme content. We also plan to study if other forms of dynamic information, such as user passivity and adoption rate, can contribute to the meme ranking process. Because only user actions are taken into account when measuring user dynamics, Furthermore, we will study if including contextual information as well as behavioral and cognitive aspects of internet users may more reliably predict meme popularity.

REFERENCE

[1] *Adithya Avvaru, and SanathVobilisetty (2020) Using BERT to Analyse meme emotions, Proceedings of the 14th International Workshop on Semantic Evaluation, pages 1094–1099.*

[2] *Alessandro Ortis , Giovanni Maria Farinella and Sebastiano Battiato (2019) An Overview on Image Sentiment Analysis: Methods, Datasets and Current Challenges, In Proceedings of the 16th International Joint Conference on e-Business and Telecommunications (ICETE 2019), pages 290-300. DOI: 10.5220/0007909602900300*

[3] *Alessandro Ortis, Giovanni Maria Farinella, Sebastiano Battiato (2018) Visual Sentiment Analysis Based on Objective Text Description of Images, International Conference on Content-Based Multimedia Indexing (CBMI), DOI:10.1109/CBMI.2018.8516481.*

[4] *Amit Pimpalkar, JebersonRetna Raj (2020a). Evaluation of Tweets for Content Analysis Using Machine Learning Models, Proceeding of Computational Intelligence and Communication Networks, IEEE Explore, pp. 454-459. DOI: 10.1109/CICN49253.2020.9242611.*

[5] *Amit Pimpalkar, JebersonRetna Raj (2020b). Influence of Pre-processing Strategies on the Performance of ML Classifiers Exploiting TF-IDF and BOW Features, Advances in Distributed Computing and Artificial Intelligence Journal, Vol. 9 N. 2, pp 49-68, DOI: 10.14201/ADCAIJ2020924968.*

[6] *Amit Pimpalkar, JebersonRetna Raj (2021). Social Network Opinion Mining and Sentiment Analysis: Classification Approaches, Trends, Applications and Issues, Advances in Intelligent Systems and Computing, Vol. 1, pp. 755-773. DOI: 10.1007/978-981-33-6981-8_60.*

[7] *Amit Pimpalkar, JebersonRetna Raj, (2022) MBiLSTMGloVe: Embedding GloVe Knowledge into the Corpus using Multi- Layer BiLSTM Deep Learning Model for Social Media Sentiment Analysis, Expert Systems With Applications, DOI: 10.1016/j.eswa.2022.117581.*

[8] *Apeksha Aggarwal, Vibhav Sharma, Anshul Trivedi, Mayank Yadav, Chirag Agrawal, Dilbag Singh, Vipul Mishra, and Hasse`neGritli (2021) Two-Way Feature Extraction Using Sequential and Multimodal Approach for Hateful Meme Classification, Complexity, Volume 2021, Article ID 5510253, 7 pages, DOI: 10.1155/2021/5510253*

[9] *Lisa Bonheme and Marek Grzes. (2020), SESAM at SemEval-2020 Task 8: Investigating the Relationship between Image and Text in Sentiment Analysis of Memes. In Proceedings of the Fourteenth Workshop on Semantic Evaluation, pages 804–816, Barcelona, DOI: 10.18653/v1/2020.semeval-1.102*

[10] *Nayan Varma Alluri, Neeli Dheeraj Krishna (2021) Multimodal Analysis of memes for sentiment extraction arXiv:2112.11850.*

[11] *Nikil Prakash T., Aloysius A. (2021) Hybrid Approaches Based Emotion Detection in Memes Sentiment Analysis, International Journal of Engineering Research and Technology, Volume 14, Number 2, pp. 151-155.*

[12] *Nikita Jain, Vedika Gupta, Shubham Shubham, Agam Madan, Ankit Chaudhary, K. C. Santosh (2021) Understanding cartoon emotion using integrated deep neural network on large dataset, Neural Computing and Applications. DOI: 10.1007/s00521-021-06003-9.lV)(0123456789,-().volV)*

[13] *Pradyumna Gupta, Himanshu Gupta, and Aman Sinha. (2020) DSC IIT-ISM at SemEval-2020 Task 8: Bi-Fusion Techniques for Deep Meme Emotion Analysis. In Proceedings of the Fourteenth Workshop on Semantic Evaluation, pages 876–884, Barcelona. DOI: 10.18653/v1/2020.semeval-1.111.*

[14] *Raj RatnPraneshAmbesh Shekhar (2020) MemeSem: A Multi-modal Framework for Sentimental Analysis of Meme via Transfer Learning, Lifelong Learning Workshop at ICML 2020, Vienna, Austria, PMLR 108.*

[15] *Riza Velioglu (2020) Detecting Hate Speech in Memes Using Multimodal Deep Learning Approaches: Prize-winning solution to Hateful Memes Challenge, arXiv:2012.12975.*

https://doi.org/10.46335/IJIES.2022.7.8.16 **e-ISSN: 2456-3463**
Vol. 7, No. 8, 2022, PP. 89-96

International Journal of Innovations in Engineering and Science, www.ijies.net

[16] *Roushan Kr. Giri, Vineet Chamoli, Prashant Singh, Subham Singh, Subhash Chandra Gupta. (2021). Using Natural Language Processing (Nlp) And Deep Learning To Analyze Sentiment In Memes. International Journal of Modern Agriculture, 10(2), 4076 – 4083.*

[17] *RupamBawankule, Amit Pimpalkar (2014), Sentence Level Text Extraction using Hierarchical Fuzzy Relational Clustering Algorithm, International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, Vol. 3 Issue 11, pp 91-95.*

[18] *ShardulSuryawanshi, Bharathi Raja Chakravarthi, MihaelArcan, Paul Buitelaar (2020) Multimodal Meme Dataset (MultiOFF) for Identifying Offensive Content in Image and Text, Proceedings of the Second Workshop on Trolling, Aggression and Cyberbullying, pages 32–41, Language Resources and Evaluation Conference (LREC 2020),*

[19] *Shobana J., Murali M. (2021) An efficient sentiment analysis methodology based on long short‑term memory networks, Complex Intell. Syst. 7, 2485–2501. DOI: 10.1007/s40747-021-00436-4*

[20] *Siddhanth U Hegde, AdeepHande, RubaPriyadharshini, SajeethaThavareesan, and Bharathi Raja Chakravarthi (2021). Tamil Troll Meme Classification: You need to Pay more Attention. In Proceedings of the First Workshop on Speech and Language Technologies for Dravidian Languages, pages 180–186, Kyiv.*

[21] *SravaniBoinepelli, Manish Shrivastava, and Vasudeva Varma. (2020). SemEval-2020 Task 8: An Overview of Simple Text Classification Methods for Meme Analysis. In Proceedings of the Fourteenth Workshop on Semantic Evaluation, pages 1190–1194, Barcelona. DOI: 10.18653/v1/2020.semeval-1.157.*

[22] *Vishal Keswani, Sakshi Singh, Suryansh Agarwal, and Ashutosh Modi. 2020. IITK at SemEval-2020 Task 8: Unimodal and Bimodal Sentiment Analysis of Internet Memes. In Proceedings of the Fourteenth Workshop on Semantic Evaluation, pages 1135–1140, Barcelona. DOI: 10.18653/v1/2020.semeval-1.150.*

[23] *https://www.kaggle.com/parthplc/facebook-hateful-meme-dataset*

[24] *https://www.kaggle.com/williamscott701/memotion-dataset-7k*

[25] *https://www.kaggle.com/datasets/adityajn105/flickr8k*