Original Article
ISSN (Online): 2582-7472

ShodhKosh: Journal of Visual and Performing Arts
June 2024 5(6), 595–600

DATA ANALYTICS SYSTEM FOR OFFENSIVE MEMES TEXT CLASSIFICATION IN SOCIAL NETWORK

Dr. Gowsic K[1], Shanmuga Sudhan K[2] Yogeshwaran H[3], Vignesh M[4], Deena R[5]

[1]Associate professor, Department of Computer Science and Engineering, Mahendra Engineering College
[2, 3, 4, 5]UG students, Department, of Computer Science and Engineering, Mahendra Engineering College

Corresponding Author
Shanmuga Sudhan K,
ssudhan051@gmail.com

DOI
10.29121/shodhkosh.v5.i6.2024.4162

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

ABSTRACT

Sentiment analysis has evolved as a pivotal element of understanding mortal feelings and opinions in the digital period. Memes, as a popular form of online communication, frequently synopsize sentiments in a visually engaging manner. In the realm of online communication, the frequence of multi-modal content, particularly memes combining textbook and images, has raised enterprises regarding the dispersion of obnoxious or dangerous material. This exploration proposes a new approach grounded on deep literacy ways for the bracket of obnoxious memes in multi- modal data sets to address this issue. The vital task involves the comprehensive analysis of both textual and visual factors to directly identify and classify unhappy content. We're proposing a new approach in this study. sentiment analysis from meme images using Optical Character Recognition (OCR) technology with Deep literacy algorithms for textbook bracket and image bracket using Conventional neural network algorithm. Our approach involves rooting textbook content from meme images using OCR, enabling the analysis of textual rudiments to infer sentiments. Through this innovative approach, we aim to enhance analysis capabilities in the visual content using pretrained Conventional neural network model. The proposed methodology demonstrates promising results, slipping light on the eventuality for OCR driven VADER sentiment analysis and Pretrained CNN to classify the nuanced feelings bedded in meme culture and contribute to a further comprehensive understanding of online sentiment dynamics.

Keywords: Memes Classification, Convolutional Neural Network, Optical Character Recognition, VADER Algorithm

1. INTRODUCTION

The proliferation of social media users in recent years has been spurred by the ease of access to the internet. Individuals are becoming increasingly outspoken and desirous of having a large audience hear their voices. Both the person and society can be greatly impacted by a post on social media. A social media post has occasionally sparked riots and hate crimes in a community. A social media meme might personally result in despair or even suicidal thoughts. Memes have become a popular tool for people to express their opinions on social media in recent years. A meme can be a picture, a video, or just some words that is funny by definition. A meme is considered offensive if it greatly irritates, infuriates, or causes other people to feel bad.

Memes are considered offensive when they offer incisive criticism on societal concepts, current events, and cultural icons with the intention of attacking specific people or groups, such as minorities or homophobes. Political hatred can now be expressed through memes as well. Thus, memes must be effectively filtered automatically in order to prevent hostile content from spreading. Memes frequently use text overlayed on photos. Put differently, memes are multimodal in nature, and it's unpredictable whether the objectionable content is linked to the backdrop image or the embedded text. The outcomes of a single modality do not allow them to be labelled as offensive. Upon examining just, the image or word, we might not be able to discern any damaging content; but, when all modalities are taken into account, the context is altered. Thus, the goal of this study is to build a multimodal system that can process the meme's visual and text simultaneously and more accurately classify them as offensive or non-offensive. Fig 1 shows the memes datasets.

Fig 1: MEMES DATASETS IMAGES

2. RELATED WORK

Rui Cao, et.al,…[1] proposed a multimodal prompt- grounded frame called PromptHate, which prompts and leverages the implicit knowledge in the PLM to perform spiteful memeclassification.In addition to the image captions as input to the PLMs, the uprooted realities and demographic information shall be used as supplementary information. In addition to the image captions as input to the PLMs, the uprooted realities and demographic information shall be as supplementary information. Note that the contextual background information is still missing in the image caption and supplementary information, indeed though the uprooted fresh information may contain crucial information about the meme.However, similar as' Muslim', If we use reality information to identify a gormandizer in the meme and decide an acronym from it. still, the supplementary information doesn't give any concrete knowledge as to whether Muslims don't consume pork.A possible reason could be in the many- shot setting, the Prompt Hate relies more on the marker words ’ semantics to prize implicit knowledge for spiteful meme bracket. therefore, the marker words with the aligned semantic class will give better environment in the prompt to ameliorate spiteful meme bracket when there are inadequate compliances in training cases.Again, when Prompt Hate is trained with enough cases, the representations of the marker words are streamlined to be near to the spiteful meme bracket task.

Shardul Suryawanshi, et.al,…[2] To develop a systemic frame that uses the NLI task in order models that have been fine tuned for Emotion Analysis, Sentiment Analysis, Offensive Tweet Bracket and NLI tasks are farther meliorated using recently converted data. For each model, three ablations have been carried out to gain a more complete understanding of how the models bear. In addition, a detailed analysis of the task's quantitative and qualitative crimes is handed. Due to the millions of trainable parameters, all multi modal SOTAs are complex and computationally delicate. thus, by taking advantage of NLI's task, we're proposing to transfigure the problem of bracket of Multi modal Offensive Memes into a uni-modal General Text Bracket Problem in order to simplify and more fluently break this issue.

Abhinav Kumar Thakur, et.al,…[3] For the bracket of IMs, particular information is given on XPlore multimodal styles. We calculate on the general idea of Case- Grounded logic(CBR), where a vaticination can be traced back to analogous memes that the system has observed at training time. CBR is chosen because it allows us to gain a clear understanding of model logic and continues to use ultramodern OTA representation literacy tools, given the complexity of IT systems. In order to grease relative analysis of exemplifications collected from all our models in each IM, a stoner friendly interface should be created.

A neurosymbolic frame, with explicatory exemplifications and prototype prognostications, should be espoused in the case of IM bracket tasks. Both styles use a to make multimodal tasks unimodal. originally, frozen pretrained model to we transfigure the data from image- textbook- marker format into premise- thesis-prize features from meme in the transmission literacy process, with separate downstream bracket models that work these features for making final opinions. It's easy to compare the combination of the point birth model and the explanation system used because of the modular nature of the frame.

Biagio Grasso, et.al,…[4] demonstrate that KERMIT is suitable to recoup applicable contextual information from Concept Net and use it effectively in the bracket process with a view to perfecting performance. In particular, In the Facebook Hateful Memes dataset, the proposed system achieved a state of the art performance and was comparable to the most recent competitors.Overall, this work has shown that it is effective to integrate external knowledge into the bracket process and provides a roadmap for future exploration of dangerous meme discoveries in terms of how AI and knowledge discovery can play an important role in improving content moderation. To that end, our concept KERMIT uses a MANN to store a knowledge enriched information network, representing the reality of memes and their associated common sense deduced from Concept Net. In particular, the KERMIT memory block consists of a number of boxes that each contain part of the knowledge enriched information network. In addition,KERMIT uses an attention medium to automatically delete the most accurate source of information on dangerous memes, so that they can be properly classified. By taking into account contextual information and relevant data from outside sources, this approach allows our model to integrate external knowledge into the decision making process, enabling it to remove dangerous content from memes.

Yang, Chuanpeng, et.al,…[5] Proposed a scalable frame for dangerous memes discovery(denoted as ISM) is proposed by learning modality-steady and modality-specific representations via graph neural networks, which is complementary to the double- aqueduct models. In order to give a comprehensive and disaggregated view of forspanning up ISM. The experimental results show that ISM offers a stable increase over birth, and is suitable to contend against being styles for discovery of dangerous memes. nullification studies and case studies also show the acceptability and rationality of each factor. A scalable dangerous memes discovery frame ISM is proposed to learn modality-steady and modality-specific representations via graph neural networks, furnishing a comprehensive and disentangled view of memes by reducing the modality gap and aligning image- textbook couples.

3. EXISITNG METHODOLOGIES

In the realm of memes textbook bracket, Count Vectorizer and TF- IDF, two abecedarian ways in natural language processing, play vital places in rooting meaningful features from textual data When each row corresponds to the content and every column represents some word in this language, Count Vectorizer is used for converting a textbook train into a matrix representation. The number of each term on the matrix document shall indicate its frequence in circumstance. As a result, all documents are represented by vectors with values indicating the number of circumstances for each term.However, Count Vectorizer doesn't have the moxie to determine the applicability of terms beyond the frequence of use. In discrepancy, the TFIDF increases this by taking into account both the frequence of words in a document and their meaning.(Term frequence) and its oddity within the whole corpus of inverse document frequentness. In this approach, to capture their discriminational powers, advanced weights are allocated to terms which are generally set up in the document but not uncommon throughout the entire corpus. In substance, compared to Count Vectorizer, the TFIDF allows more precise understanding of memes that are particularly useful for textbook bracket tasks where identification of features is critical in order to directly classify textbooks. Graph Neural Networks (GNNs) offer a important frame for memes bracket by using the essential structure memes, the ISM system divides each of the and relationship within meme data. When modalities into two different spaces with distinctive characteristics, while learning parallels to reduce the gap between themodalities. Moreover, the mainstream multimodal binary models of aqueducts similar as CLIP, ALBEF and BLIP could be a base bumps represent rudiments similar as images, textbook, or druggies and edges represent connections or relations between these rudiments, data can be represented as a graph in memes bracket.

These detailed, interdependent connections can framework utilizes advanced machine learning be effectively modeled and exploited by GNNs techniques, such as Optical character in order to increase the effectiveness recognition and VADER algorithm to analyse ofclassification. GNNs operate by iteratively meme content and extract relevant features for adding up information from neighbouring classification. In order to find patterns and bumps in the graph to modernize each knot's associations of different meme attributes, these point representation. This process allows models are trained on the labeled dataset. so GNNs to capture both original and global that they can be accurately classified. dependences within the graph, enabling them Furthermore, to analyse the text content of to effectively capture the semantics and memes, sentiment analysis on user reactions environment of meme data.

4. PROPOSED METHODOLOGIES

The proposed meme classification system use linguistic processing techniques. These comprises several crucial components aimed at analyses contribute to a comprehensive analyzing and categorizing memes effectively. understanding of the meme content, allowing The process begins with the collection of a for more complex classification decisions.

Diverse dataset containing labeled memes, Furthermore, the classification system may ensuring a comprehensive representation of include mechanisms for continuous learning various content types, styles, and sources. This and adaptation, allowing it to evolve over time dataset serves as the foundation for training as new memes emerge and user preferences and evaluating the classification model. Once change. This involves periodically retraining the dataset is assembled, the next step involves the classification model with updated data to implementing the classification framework to ensure its effectiveness in classifying memes in classify memes within social networks. This the ever-evolving landscape of social media. entails deploying the trained model to Overall, the proposed meme classification automatically categorize memes shared across system integrates data collection, model different social media platforms. The system development, and real-time classification may leverage application programming capabilities to provide a comprehensive interfaces (APIs) provided by social networks to solution for analysing and categorizing memes access meme content and perform in social networks. See fig 2 for the proposed classification in real-time. The classification framework.

ALGORITHM STEPS

- Optical Character Recognition: -
- Extraction of characters' boundaries from image,
- Building a Bag of Visual Words (BOW) framework in remembering the Character images,
- Loading trained Model,
- Consolidating predictions of characters.

TEXT MINING ALGORITHM

ALGORITHM STEPS:

Collecting memes text and performing preprocessing steps to remove noisy words is a first step.

- **TOKENIZATION:** The given document is False positive (FP): despite the absence of text in the sample, a discovery system gives an accurate result to the sample.

True negative (TN): indeed though the sample doesn't contain any text, a positive test result is generated when styles are set up. False negative (FN): In malignancy of the fact that The contains the malignancy of the fact that the sample contains the text, he discovery system results in a positive test result for the sample.

considered as a string and relating single word Precision = P (1)

in document i.e., the given document string is divided into one unit or token $P+P$

That Fp is equal is zero. If Framework

- **JUNKING OF STOP**: Word In this step the junking Programme is increased. There will be a of usual words like a, an, but, and, of, the etc . reduction in perfection and An increase in is done denominators that lead to differing results From
- **STEMMING:** A stem is a combination of words what we want. which have the same or analogous meaning.

This system describes the base of the given Recall= (2)

word. Inflectional and derivational stemming are two types of system.

VADER ALGORITHM

- Import demanded library $P+N$
- A good classifier should have a recall of one(high).
- Take the text for which it is to be calculated In order to infer the FN is zero. Only if the
- when calculating opposition and intensity. denominator and numerator are identical as in

- Calculate the VADER score using the for TP= FN does one recall an equal number. circle.
  VADER returns four factors that are With the addition of FNs, the recall value
- favourable, unfavourable and indifferent. decreases which is undesirable and the smallest
  Determine whether the text is favourable,common value increases.Therefore, unfavourable or indifferent according to a consideration should be given to delicacy and conflation score recall in the Statistical Measures. The F1 Score, still, If the opposition is> 0 text is considered the computation to take into account the as favourable.

5. EXPERIMENTAL RESULTS

We'll be suitable to input meme images into this study, the text of the prize and determine whether or not they are obnoxious. In the assessment of performance, account shall also rates.

TRUE POSITIVE (TP): A positive opinion of the sample is given by the Seeing System, and this text appears in the sample rates.

FALSE POSITIVE (FP): despite the absence of text in the sample, a discovery system gives an delicacy and recall.

F1 Score =2* $Prcso*ca$

$Prcso+ca$

ALGORITHM	F1 SCORE	
Naives Bayes	60%	
Random forest	70%	
Text mining with VADERAlgorithm	90%	

accurate result to the sample.

TRUE NEGATIVE (TN): indeed though the sample doesn't contain any text, a positive test result is

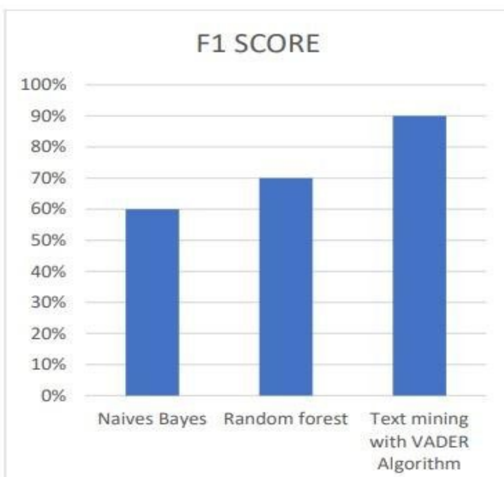

Fig 3: Performance chart

The proposed classification of Memes, based on the performance graph, has been shown to be more efficient in comparison with current algorithms.

6. CONCLUSION

In conclusion, the integration of Optical Character Recognition (OCR), text mining, and VADER sentiment analysis represents a potent approach to memes classification, offering a multi- faceted solution to understanding and categorizing meme content. Using OCR, the system extracts text from meme images, providing valuable insight into the textual components that often accompany visual elements.This enables a deeper analysis of meme content, capturing nuances and subtleties that may not be immediately apparent from the images alone. Furthermore, text mining techniques allow for the extraction of key features and patterns from the textual content of memes. Through processes such as tokenization, stemming, and topic modelling, the system can uncover underlying themes, sentiments, and linguistic characteristics present in meme captions or overlays. This textual analysis enriches the classification process, providing additional context and understanding of meme content beyond its visual representation. By analysing sentiment scores, the system can discern whether meme captions or text overlays convey positive, negative, or neutral sentiments, further enhancing the classification process. This generated when styles are set up.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

Cao, Rui, et al. "Prompting for multimodal hateful meme classification." arXiv preprint arXiv:2302.04156 (2023).

Suryawanshi, Shardul, eal."Multimodal Offensive Meme Classification with Natural Language Inference." Proceedings of the 4th Conference on Language, Data and Knowledge.2023.

Thakur, Abhinav Kumar, et al. "Explainable Classification of Internet Memes." (2023).

Grasso, Biagio, et al. "KERMIT: Knowledge- EmpoweRed Model In harmful meme deTection." Information Fusion 106(2024): 102269.

Yang, Chuanpeng, et al. "Invariant Meets Specific: A Scalable Harmful Memes Detection Framework." Proceedings of the31st ACM International Conference on Multimedia. 2023.

Pramanick, Shraman, et al. "Detecting harmful memes and their targets." arXiv preprint arXiv:2110.00413 (2021).

Zhong, Xiayu. "Classification of Multimodal Hate Speech--The Winning Solution of Hateful Memes Challenge." arXiv preprint arXiv:2012.01002 (2020).

Suryawanshi, Shardul, et al."A dataset fortroll classification of TamilMemes." Proceedings of the WILDRE5–5th workshop on indian language data: resources and evaluation. 2020.

Ling, Chen, et al. "Dissecting the meme magic: Understanding indicators of virality in image memes." Proceedings of the ACM on human-computer interaction 5.CSCW1 (2021):1-24.

Kumar, Gokul Karthik, and Karthik Nandakumar. "Hate-CLIPper:

Multimodal hateful meme classification based on crossmodal interaction of CLIP features." arXiv preprint arXiv:2210.05916 (2022)