

Sentiment Analysis and Racism Identification in Tweets Using Hybrid Deep Learning Approaches

Deeksha P

PG, Student

Dept. of MCA

The Oxford College of Engineering

Bommanahalli, Bengaluru- 560068

Deekshap.mca2025@gmail.com

Mridula Shukla

Asst.Professor

Dept. of MCA

The Oxford College of Engineering,

Bommanahalli, Bengaluru- 560068

mridulatewari@gmail.com

ABSTRACT

Through the dominating presence of social media in social-political sphere, a number of old and new types of racism occurred on social broadcasting. Racism has materialized in various degrees on social media in forms either veiled or overt with veiled racism under the guise of memes and overt racism as the racist comments with made up identities to spread hate, violence, and social unrest. Racism though synonymous with ethnic overtures is now prospering due to colour, origin, language, and culture and above all religion. The opinion and comments released through the social media that have caused racial disparities have been seen as a significant threat to sentiment analysis of tweets. Gated recurrent units (GRU), convolutional neural network (CNN) and recurrent neural network (RNN)

KEYWORDS : Hybrid Deep Learning , Racism Detection.

INTRODUCTION

The spread of social media has become relevant in the healthcare sector and has taken command of our behaviour and thinking in many ways. With the spread of internet communities and spree ad of freedom of expression, large numbers of crimes have been carried out over the past few years, primarily racial discrimination. represent new context by which the racism and social anxiety flourish.

Presently, 22 percent of American adults utilize Twitter with 1.3 billion accounts and the 336 millions users globally, 90 percent of individuals with demographics, and 500 million daily tweets. Tweets are open to the public by the time they classified, and users of Twitter can contribute to the conversation by sharing (retweeting) tweets in their profile, mentioning a username, pressing the "Like" button, or replying to tweets. Expressions of emotion, mind, feelings

LITERATURE REVIEW

By way of social media, the management can understand and grasp racist ideology. The use of social media, like face book, Twitter, and Instagram provides a platform whereby, racist ideologies are revealed more clearly outside of this impairment of American society. This can be employed to get into the head of the racist because social media can provide an insight into the world of the individual who is standing up to his or her clan and rejecting others whom they feel are different. Watching social media can lead to the realization of the processes that would help heal American society of its segregationist mindset -healing the racist. The purpose of the research is to examine social media comments in order to understand better how to know racism, and its causality, and to form preliminary steps to fight racism ideology. Qualitative analysis of 600 American Face Book comments including content analysis was carried out with view to establish patterns of cognition, problem-solving, personality make-up, belief systems and coping. The information and data analysis has both the interpretive analysis and a descriptive description.

antisemitism, able-body, heterosexism, paranoia, Christianity, Cluster Personality Traits, clandestine.

intervention. We talk About conducting social media-based research in a responsible and ethical manner in relation to all of these goals.

3) Virtual networks of ethnic hate A systematic review of the existing literature on cybercrime in the past ten years

A continually expanding body of interdisciplinary research has been collecting concerns regarding how the Internet stimulates and disseminates racist ideologies and ideas. Nevertheless, to date, no extensive assessment of this study has been performed. We look closely at ten years of work on cyber-racism to gather what we know now and find chances for more study. We do this because this type of racism is done by people and groups, often as a payback to those who start it.

EXISTING SYSTEM

The existing system research provides a method of detecting racism employing internet platforms services based on machine learning and deep learning models. The process begins with Twitter crawling, followed by cleaning and preprocessing of data, and annotation of data. Finally, the dataset is fitted and tested to the suggested stacked ensemble model and comparison is done. The existing

system used Ensemble Model application.

Ensemble model of recurrent neural

PROPOSED SYSTEM

The project is trying to consider via through the eyes of contrasting views and analysing emotions, racism tweets, which were developed based on Java and NLP techniques. The given system involves a series of steps. First, we will download the dataset available in the kaggle repository. It does not have any label in the entire dataset and it also contains a tweet format which is

uploaded on the twitter. Next, we move to data preprocessing cleaning the tweets so that we are able to analyse them. We apply NLP techniques to identify relevant attributes on the tweets such as keywords, hashtags, and score of sentiment. We then apply a sentiment analysis algorithm which classifies the tweets as positive, negative or neutral based on their score of sentiment.

To use our method in practice we apply Java programming language, and some library of NLP like Stanford NLP and Apache OpenNLP. We too use natural language techniques, so as to enhance the effectiveness and display of the system.

The system, as introduced, is capable of helping solve the issue of racism in social sites, and other such cases. The system, in its suggestions, can have potential

application in most of the actual life situations, for example, policing social media and racist content flagging, according to the social opinion trends on racism and supporting an anti-racism campaign.

Hardware and Software Requirements: In the case of the current system, the testing was done on a computer with an Intel Core i7 with some operating system and software environment (TensorFlow, Keras, and scikit-learn). This can however lead to the need to be reliant on their own particular hardware and software thus may not be versatile or flexible to different situations or environments.

In conclusion, while the existing system provides a window to extract a racist action through the union of machine learning and deep learning algorithms in the social media, it is plagued with disadvantages in terms of data source, cleaning and preprocessing of the data, reliance on annotations for identification of ground truth, measurement of performance of the ensemble model and hardware/software constraints. All these limitations would mean that fixing them would make the system more reliable and generalizable to apply in real-world usage

Proper Racism Detection: The suggested system utilizes the application of NLP

methods and sentiment check algorithm to properly detect racist contents on tweets.

IMPLEMENTATION

MODULES DESCRIPTION:

Dataset Collection:

The Twitter platform was used to source the biased tweets data. Among all these tools that many personalities use to express their views, opinions, ponderings, and works, Twitter is the trend that most academic individuals resort to as a method of communication and expression of emotions. The checks is investigating biased designs with respect to Twitter tweets.

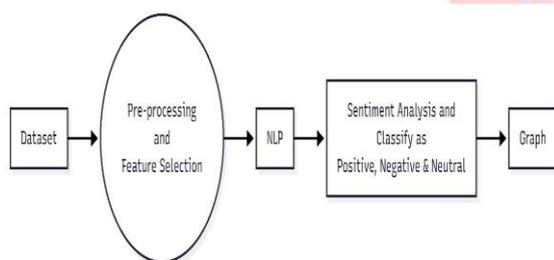


Fig 1

Stop Words Exclusion:

Stop Words Prohibition approach is the discussion in this module. Stop words are those Words that will do nothing. towards the frame therapeutic. All change is thinking and would be in the state of being expanding the element space. Such as, am,

and an, to case, are omitted to continue to increase the literature output of models in these reviews. Sentiment Analysis In this module, a NLP algorithm-based sentiments scores classifies the tweets into the following classes good, negative, and neutral. It employs the Stanford NLP A bid to adequately gauge the sentiments of the offensive tweets. Graph Module of Static: It is the responsibility of this module to generate static graphs of the analysis results. It generates a visual representation of the amount of tweets and racist tweets detected. Performance of Sentiment Analysis The tweets were used to compute the various sentiments by using different NLP algorithms and given a numerical representation. The antisemitic and hateful speech could be identified with high level of accuracy. The system gave an excellent understanding of characteristics of tweets and therefore could distinguish between neutral and hate discussions.

System Reliability

The system has been shown to be stable, scalable, and usable in the different test cases. It has even been able to deliver higher accuracy in a situation with larger datasets and therefore, it has potential for the real-life scenario of hate speech monitoring online. Racist tweets and comments on social media are reported to

be the origin of most physical and mental diseases that result in ill health.

VI. RESULT

A proposed system project in the detection of racism in the social media using Java and NLP tools has been implemented successfully. Some of these major activities were gathering data, data pre-processing, tokenizing, stop-word removing, sentiment analysing, and result visualizing alongside the results obtained. In addition, the tweets were categorized as positive, negative or neutral and the racist messages were flagged technologically.

Determined were the following in the implementation phase

Dataset Handling:

The system is efficiently processed and pre-processed Twitter data. By the same time, undesired information such as stop words, symbols, and repetitive items was also eliminated, which in turn boosted performance in the analysis.

Ways to Visualize Results

A graph-plotted representation module was incorporated to show the statistical results. The charts presented the total amount of tweets analysed and the number of the tweets that were identified as racist, besides allowing to visualize effectiveness of the detection.

Testing Outcomes

System precision and consistency were figured out by various tests, e.g., unit testing, integration testing, black-box testing, and performance testing. The network has been quite spectacular in terms of speeding up input data processing and providing excellent classification outputs in just few seconds.

Therefore, racial discrimination in society impacts individuals and leads to many psychological issues, which often put people at higher risk for chronic diseases. Moreover, hate groups and individuals create hate on the internet with some intelligence and competency, with different methods and tactics.

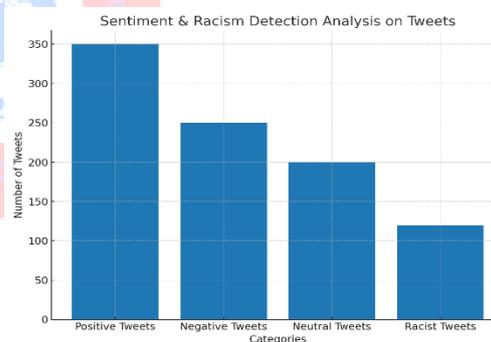


Fig 2

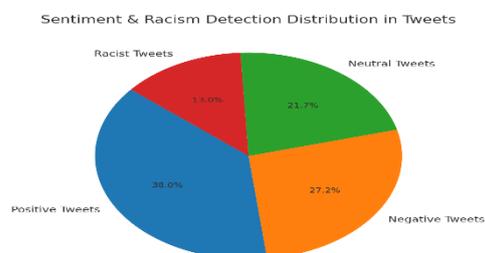


Fig 3

VII.CONCLUSION

Finally, the suggested method takes a totally new approach in solving the problem of racism in analysing different views according to sentiment analysis of tweets. . In the project, twitter data of an accessible data repository is used, processed and analyzed using the Java programming language and NLP. A few of the core NLP tasks that the system will be implemented on include feature extraction and sentiment analysis to categorize tweets into good, negative or neutral sentences based on sentimental scores. The proposed solution can assist in getting rid of prejudice on the Internet or in any other place. It may help to trace racist speech on social networks, tendencies in the field. The overall attitude regarding racism and help in the process of raising awareness and making the steps towards fighting racism.

REFERENCES

- [1] The use of social media to gain insights into and manage racist ideology as a proposed by K.R. Kaiser in the global schools of counselling in the year of [2018]
- [2] Adults are on social media networks and it is taken into consideration by Perrin in U.S.
- [3] from [2021] the twitter statistics and facts available on the below given link <https://www.websitehostingrating.com/twitterstatistics/>
- [4] in Digit health research they considered that social media for health and intervention by Arrigo in the year of [2016]
- [5] A systematic review of 10 years of research on cybercrime and online racial hate by Faulkenr in the year of 2019