IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 12, April 2025
Impact Factor: 7.67

Unmasking Online Hate: Sentiment Based Racism Detection in Tweets

[1]Shivam Pandav, [2]Shreyash Jadhav, [3]Sahil Kapadnis
Student, Department of Computer Engineering[1,2,3]
Yadavrao Tasgaonkar Institute of Engineering and Technology, Bhivpuri,
[1] pandavshivam5135@gmail.com [2] toshreyashjadhav2@gmail.com, [3] sahil.kapadnis@gmail.com

Abstract: *The Internet's widespread availability has drastically altered how we view the world. One of the most significant data sources for academics is Twitter, which is presently one of the top platforms among the several extant social networks. Social media may be utilized as real-world sensors to gauge the pulse of cultures. However, the vast and unfiltered stream of social media posts today raises societal concerns, particularly when these posts contain racism directed at a particular person or group. Social media, particularly Twitter, has been utilized in recent years to propagate anti-Racist messages. Governments and non-governmental organisations (NGOs) are concerned about the potential adverse effects that these messages may have on people or on society under this situation. In this study, we suggest Stanford NLP's Sentiment Analysis of Tweets for Racism Detection. The project uses Stanford NLP sentiment analysis to look for racism in tweets. The initial input is the Twitter dataset. Once the text data has been pre-processing, sentiment analysis is used to categorize tweets as racist or not racist. Through automatic moderation, the findings will be utilized to increase awareness and stop the propagation of prejudice online. Our key contribution is the use of Stanford NLP to obtain promising outcomes in area of racism*

Keywords: Racism, social media, Twitter, NLP.

I. INTRODUCTION

In today's digital landscape, platforms like Twitter have become breeding grounds for the spread of racism and hate speech, often masked in subtle, coded language. This project, titled "Unmasking Online Hate: Sentiment-Based Racism Detection in Tweets Using Java and NLP Techniques," aims to tackle this issue by developing a system that leverages Natural Language Processing (NLP), sentiment analysis, and data mining to detect, monitor, and mitigate racist content in tweets. The system focuses on identifying discriminatory language related to ethnicity, religion, culture, and language, while also analyzing the emotional and psychological impacts of such content. Despite challenges like processing massive volumes of real-time data, recognizing cross-cultural expressions, and balancing content moderation with free speech, the project aspires to promote online inclusivity and mental well-being. Future directions include enhancing feature extraction, reducing algorithmic bias, enabling real-time alerts, improving user interfaces, and deploying the system in real-world scenarios to ensure its effectiveness and scalability in combating racism on social media.

II. LITERATURE REVIEW

1) K. R. Kaiser, D. M. Kaiser, R. M. Kaiser, and A. M. Rackham
"Using social media to understand and guide the treatment of racist ideology" Published Year: vol. 8, Apr. 2018.
Social media, including sites such as Face Book, Twitter and Instagram, provides a platform for racist ideology, making this dysfunction of American society more evident. Social media can provide insight into the world of the racist – individuals who cling to their tribal identities, irrationally rejecting those who they perceive as different. Studying social media may provide insight into processes that can assist in healing American society of its segregationist views – a way toward healing the racist. The purpose of this paper is to analyze social media posts to better understand racism, its causality, and to develop initial steps for addressing racist ideology. A qualitative review consisting of content analysis of 600 American Face Book posts was completed to reveal patterns in cognition, problem solving, personality

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 12, April 2025
Impact Factor: 7.67

structures, belief systems, and coping styles. The content analysis consists of both a descriptive account of the data and an interpretive analysis. Keywords: Racism, social media, violence, social conditioning, sexism, ageism, anti-Semitism, able-bodyism, heterosexism, paranoia, Christianity, Cluster B Personality Traits, clandestine.

2) D. Arigo, S. Pagoto, L. Carter-Harris, S. E. Lillie, and C. Nebeker
"Using social media for health research: Methodological and ethical considerations for recruitment and intervention delivery"
Published Year: vol. 4, Jan. 2018
As the popularity and diversity of social media platforms increases so does their utility for health research. Using social media for recruitment into clinical studies and/or delivering health behavior interventions may increase reach to a broader audience. However, evidence supporting the efficacy of these approaches is limited, and key questions remain with respect to optimal benchmarks, intervention development and methodology, participant engagement, informed consent, privacy, and data management. Little methodological guidance is available to researchers interested in using social media for health research. In this Tutorial, we summarize the content of the 2017 Society for Behavioral Medicine Pre-Conference Course entitled 'Using Social Media for Research,' at which the authors presented their experiences with methodological and ethical issues relating to social media-enabled research recruitment and intervention delivery. We identify common pitfalls and provide recommendations for recruitment and intervention via social media. We also discuss the ethical and responsible conduct of research using social media for each of these purposes.

3) A.-M. Bliuc, N. Faulkner, A. Jakubowicz, and C. McGarty
"Online networks of racial hate: A systematic review of 10 years of research on cyberracism" Published Year: vol. 87, Oct. 2018.
The ways in which the Internet can facilitate the expression and spread of racist views and ideologies have been the subject of a growing body of research across disciplines. To date, however, there has been no systematic reviews of this research. To synthesise current knowledge on the topic and identify directions for future research, we systematically review a decade of research on cyber-racism as perpetrated by groups and individuals (i.e., according to the source of cyber-racism). Overall, the cyber- racism research reviewed shows that racist groups and individuals use different communication channels, are driven by different goals, adopt different strategies, and the effects of their communication are distinctive. Despite these differences, both groups and individuals share a high level of skill and sophistication when expressing cyber-racism. Most of the studies reviewed relied on qualitative analyses of online textual data. Our review suggests there is a need for researchers to employ a broader array of methods, devote more attention to targets' perspectives, and extend their focus by exploring issues such as the roles of Internet in mobilising isolated racist individuals and in enabling ideological clustering of supporters of racist ideologies.

4) M. A. Price, J. R. Weisz, S. McKetta, N. L. Hollinsaid, M. R. Lattanner, A. E. Reid, and M. L
"Meta-analysis: Are psychotherapies less effective for black youth in communities with higher levels of anti-black racism?"
Published Year: vol. 71, Feb. 2021
Method: A subset of studies from a previous meta-analysis of 5 decades of youth psychotherapy randomized controlled trials was analyzed. Studies were published in English between 1963 and 2017 and identified through a systematic search. The 194 studies (N = 14,081 participants; age range, 2-19) across 34 states comprised 2,678 effect sizes (ESs) measuring mental health problems (eg, depression) targeted by interventions. Anti-Black cultural racism was operationalized using a composite index of 31 items measuring explicit racial attitudes (obtained from publicly available sources, eg, General Social Survey) aggregated to the state level and linked to the meta-analytic database. Analyses were conducted with samples of majority-Black (ie, ≥50% Black) (n = 36 studies) and majority-White (n = 158 studies) youth.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 12, April 2025
Impact Factor: 7.67

5) D. Williams and L. Cooper
"Reducing racial inequities in health: Using what we already know to take action" Published Year: vol. 16, no. 4, p. 606, Feb. 2019.
This paper provides an overview of the scientific evidence pointing to critically needed steps to reduce racial inequities in health. First, it argues that communities of opportunity should be developed to minimize some of the adverse impacts of systemic racism. These are communities that provide early childhood development resources, implement policies to reduce childhood poverty, provide work and income support opportunities for adults, and ensure healthy housing and neighborhood conditions. Second, the healthcare system needs new emphases on ensuring access to high quality care for all, strengthening preventive health care approaches, addressing patients' social needs as part of healthcare delivery, and diversifying the healthcare work force to more closely reflect the demographic composition of the patient population. Finally, new research is needed to identify the optimal strategies to build political will and support to address social inequities in health. This will include initiatives to raise awareness levels of the pervasiveness of inequities in health, build empathy and support for addressing inequities, enhance the capacity of individuals and communities to actively participate in intervention efforts and implement large scale efforts to reduce racial prejudice, ideologies, and stereotypes in the larger culture that undergird policy preferences that initiate and sustain inequities.

III. METHODOLOGY

Dataset Collection:
The racism tweets dataset is collected from Twitter. Twitter has been the first choice of the majority of researchers for text and sentiment analysis due to its being the most common platform widely used by a large number of people to express their feelings, views, comments, and opinions. In particular, this study intends to study the racism trends based on Twitter posts. For data collection, tweets related to racist comments have been collected. This module involves collecting the dataset from a publicly available repository, such as Kaggle, which contains tweets posted on Twitter. The collected data may not have any labels and would require further processing for analysis.
Input Dataset:
In this module we develop our proposed system to accept the input dataset which is referred from kaggle. This dataset doesn't contain any label. It takes the tweets as input and uses natural language processing techniques to identify the racism tweets by analyzing it.
Data Tokenization:
This module focuses on cleaning and preparing the collected data for analysis. It may involve tasks such as removing irrelevant information, handling missing data, and normalizing the text data. Data preprocessing is crucial for ensuring the accuracy and reliability of the subsequent analysis. In this module, we make the Data Tokenization process. Tokenization is the process of splitting natural texts into tokens without any white spaces. It involves breaking sentences down into constituent words set. Although looks simpler and straightforward, deciding which tokens are appropriate is not a trivial task.
Stop Words Exclusion:
In this module we perform the Stop Words Exclusion process. Stop words are words that do not contribute to the training of the system. Instead, they create complexity by increasing the feature space. So, stop words such as a, am, and an, etc., are removed to increase the learning efficiency of models in this study.
BoW:
The BoW is another commonly used feature extraction used in NLP tasks. It is the most convenient and adaptable approach to get a document's features. The Word's histogram within the text is examined in BoW. The frequency of the words is employed as a function for the training of the set. The BoW approach is implemented in this study by utilizing the Count Vectorizer. The technique of obtaining numerical vectors by transforming a textual data set is termed vectorization. The frequency of words is counted indicating that tokens have been counted and making the token vectors. The BoW assigns a value to every attribute based on the frequency of those features.
Sentiment Analysis:

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 12, April 2025
Impact Factor: 7.67

This module involves applying a sentiment analysis NLP algorithm to classify the tweets as positive, negative, or neutral based on their sentiment scores. It uses NLP Stanford model to accurately determine the sentiment expressed in the tweets related to racism.

Static Graph Module:

This module is responsible for plotting static graphs with the results obtained from the analysis. It generates a visual representation of the total tweets and racism tweets identified from it.

IV. SYSTEM ARCHITECTURE

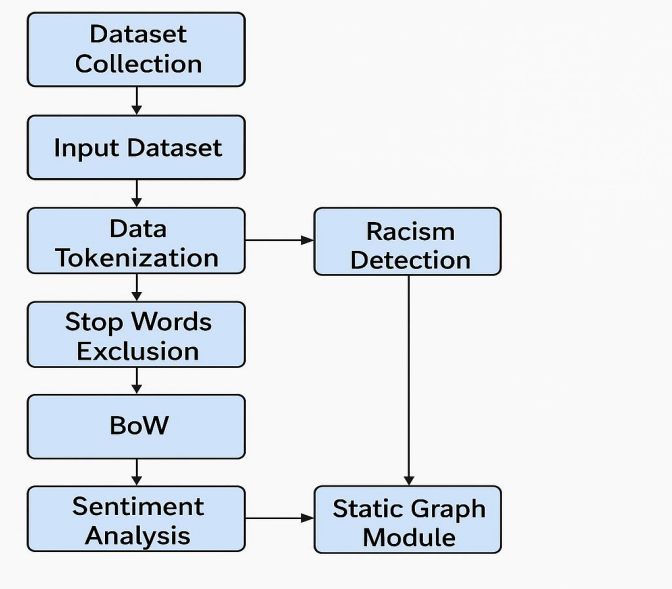

Fig. Class Diagram

ADMIN
Login
Upload data () Preprocessing () NLP Analysis () Racism Detection() Results () Graph Analysis()

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 12, April 2025
Impact Factor: 7.67

V. RESULT

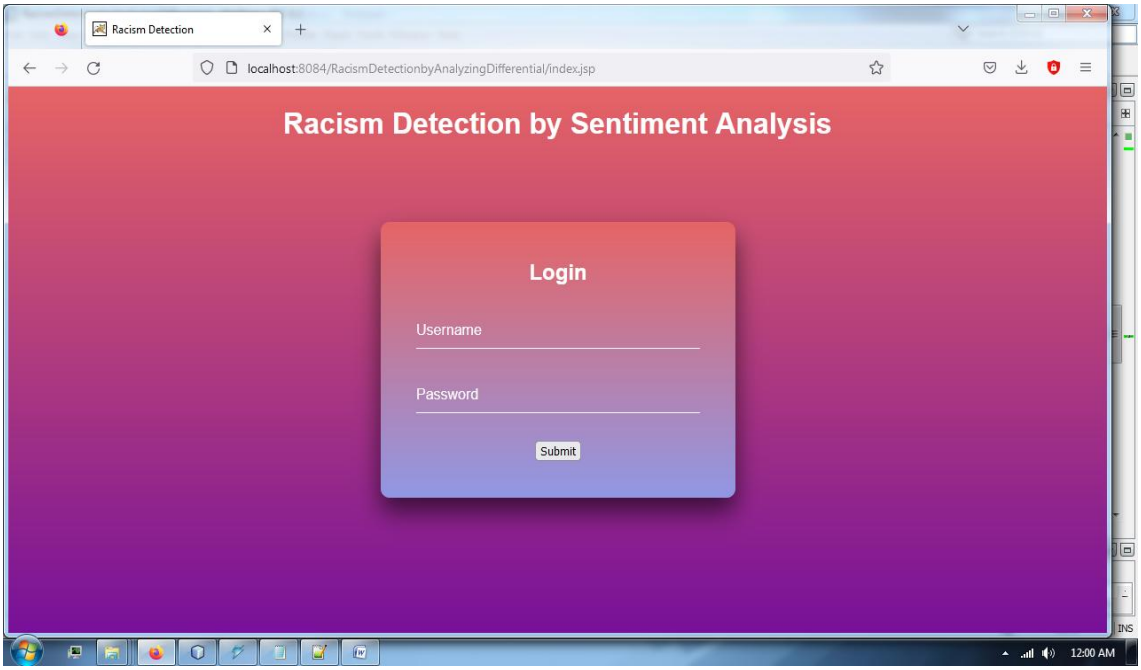

Fig 1. Login Page

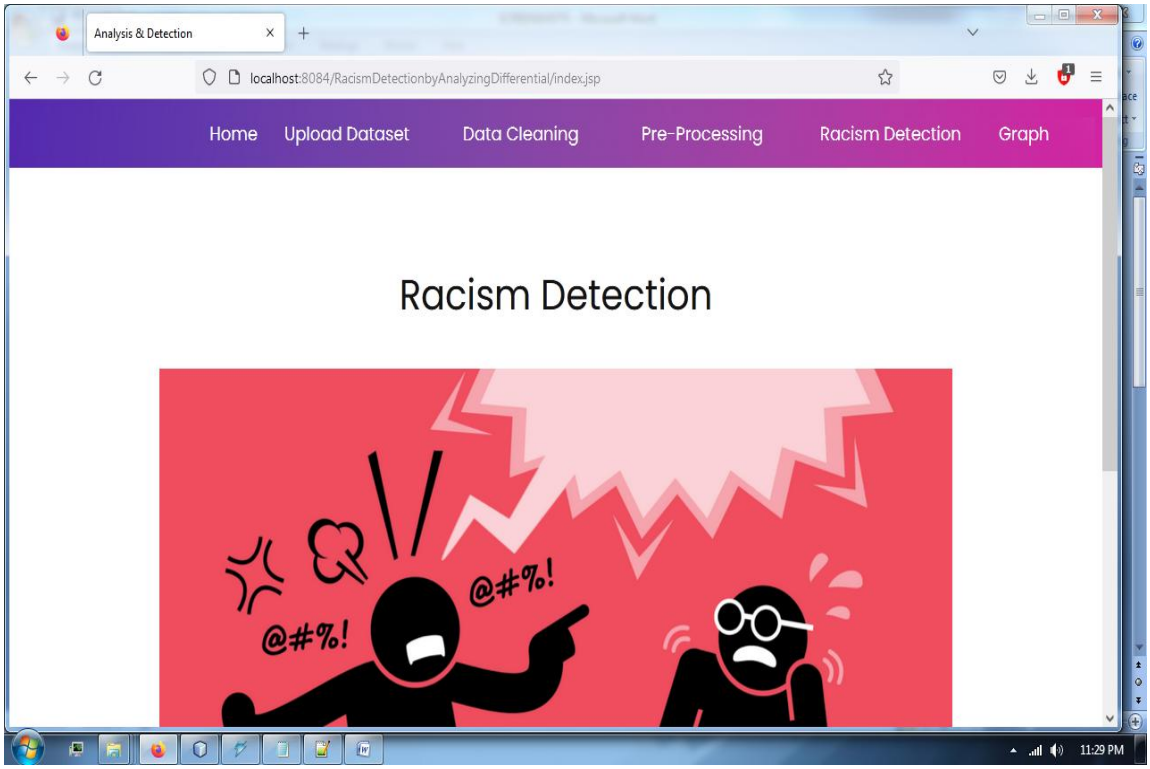

Fig 2. Home Page

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 12, April 2025
Impact Factor: 7.67

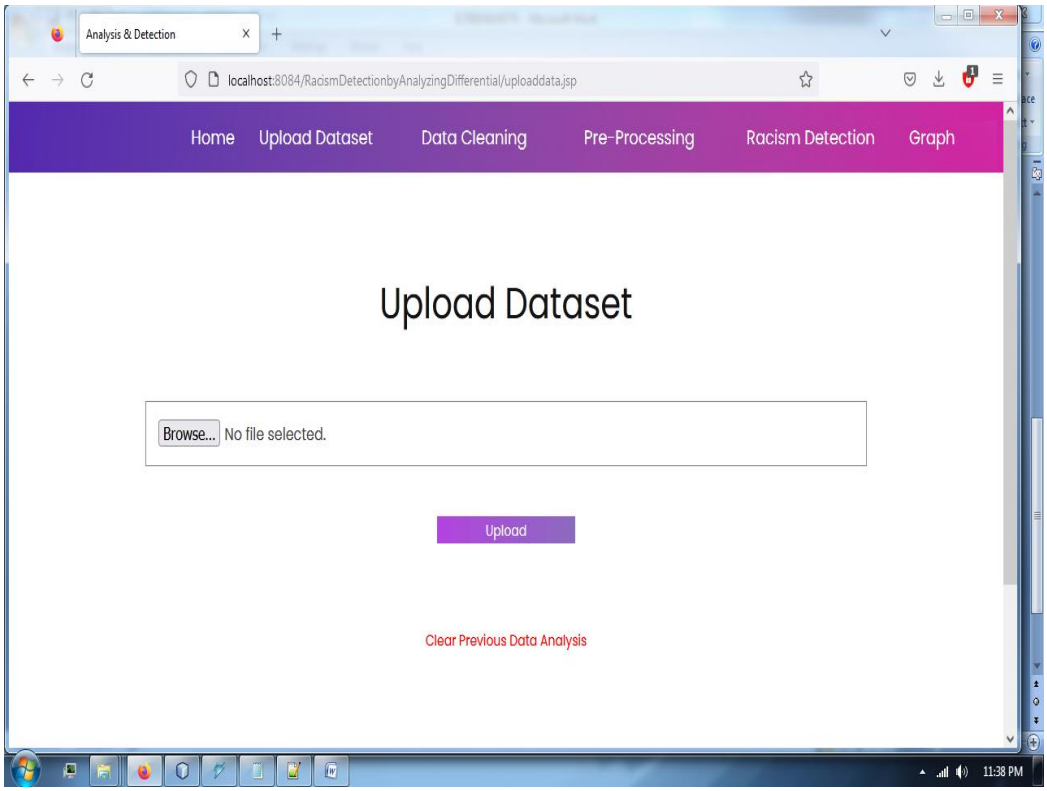

Fig 3. Upload Dataset

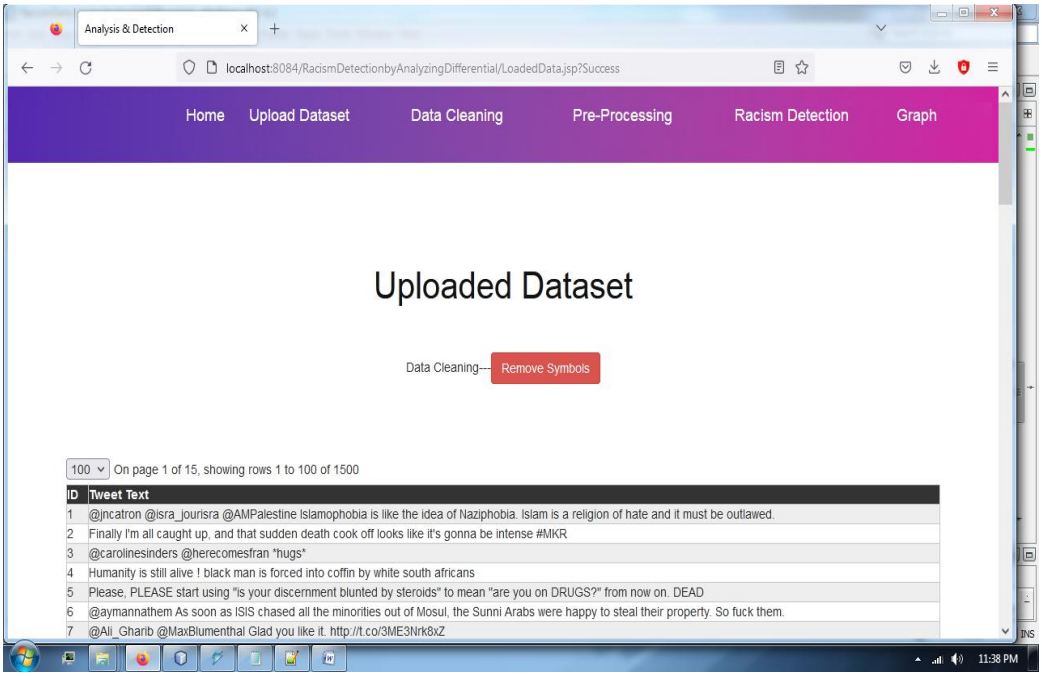

Fig 4. Uploaded Dataset

DOI: 10.48175/IJARSCT-25924

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 12, April 2025
Impact Factor: 7.67

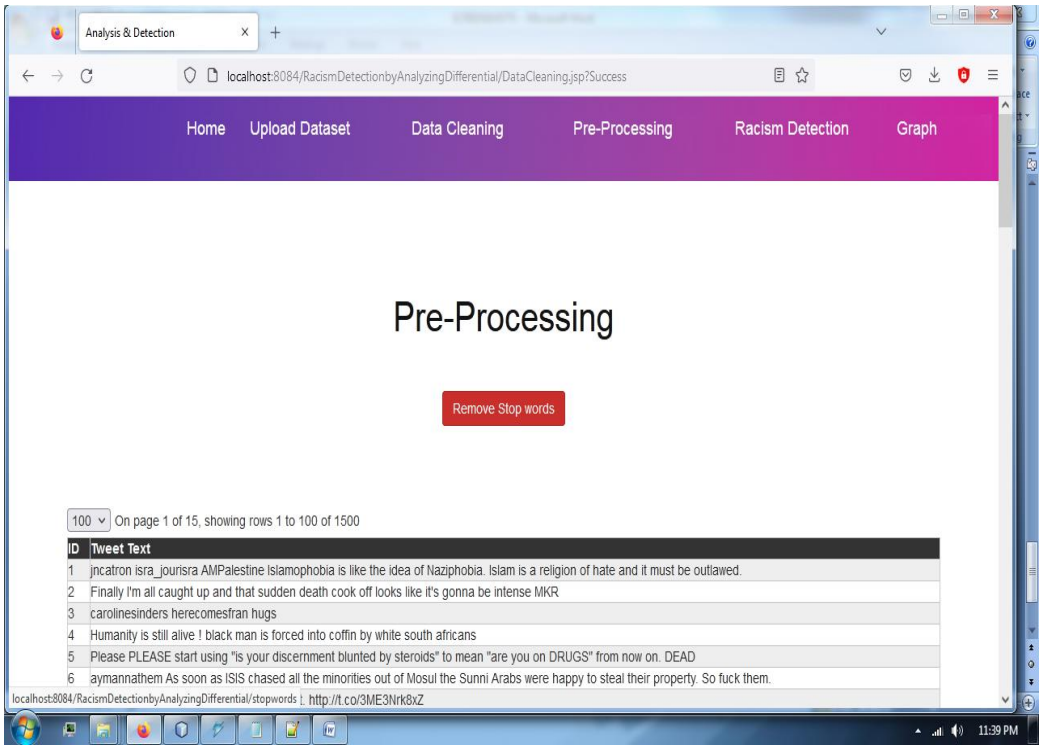

Fig 5. Data Pre-processing

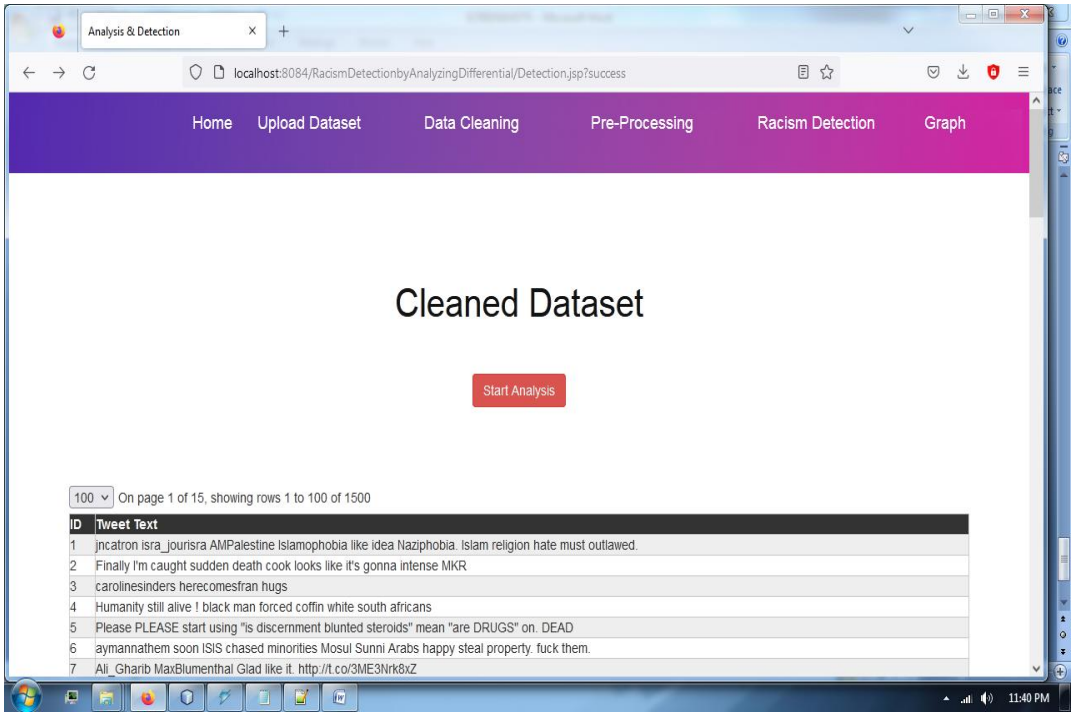

Fig 5. Cleaned Dataset

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 12, April 2025
Impact Factor: 7.67

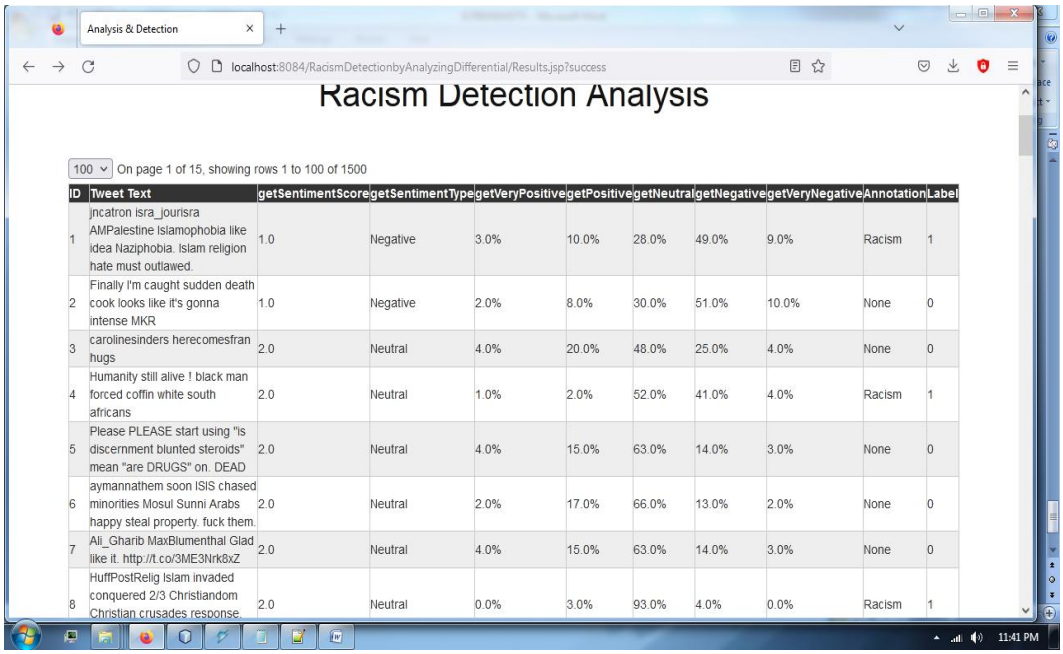

Fig 6. Analysis Page (Result)

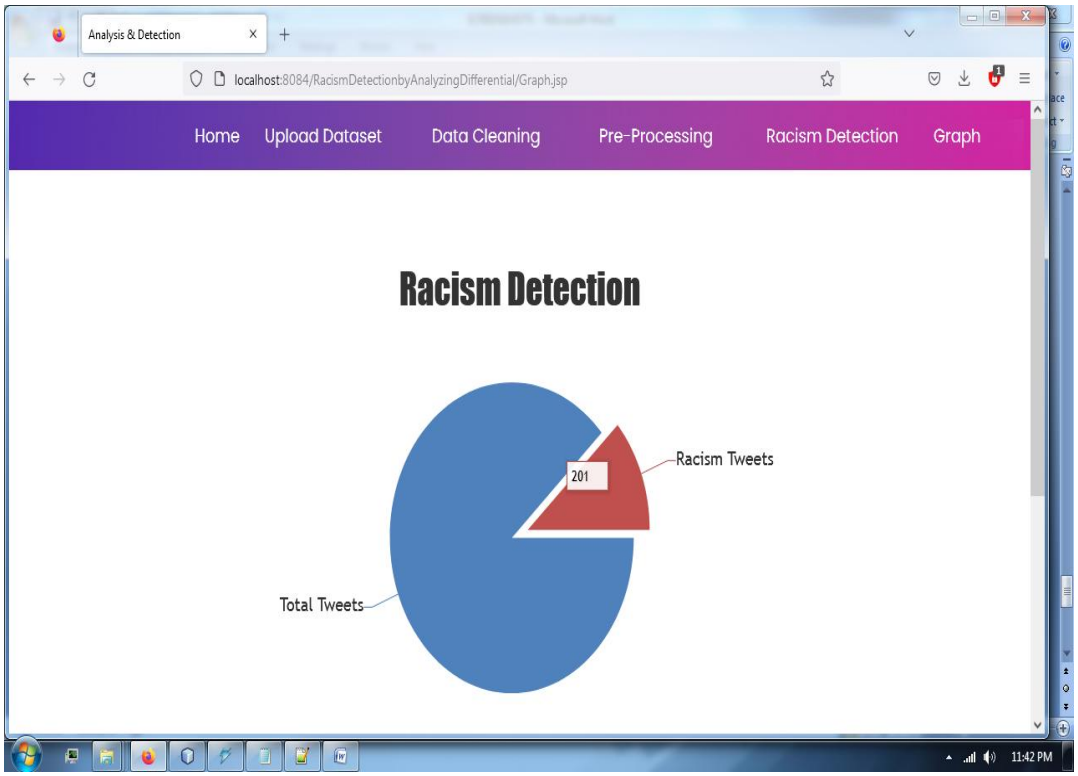

Fig 7. Final Output

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 12, April 2025
Impact Factor: 7.67

VI. CONCLUSION

In conclusion, the proposed system presents an innovative approach for addressing the issue of racism by analyzing differential opinions through sentiment analysis of tweets. The project utilizes Java programming language and NLP techniques to collect, preprocess, and analyze tweet data from a publicly available repository. The system incorporates various NLP tasks such as feature extraction and sentiment analysis to classify tweets as positive, negative, or neutral based on their sentiment scores. The proposed system has the potential to contribute to combating racism in social media and beyond. By automatically detecting racist content in tweets, it can help in monitoring social media for discriminatory language, identifying trends in public opinion on racism, and supporting efforts to raise awareness and take necessary actions to combat racism. Moreover, the system can be extended to other real-world applications such as identifying hate speech, bias, and discrimination in other forms of text data, and supporting decision-making in organizations working towards promoting diversity and inclusion. Overall, the proposed project presents a valuable contribution to the field of NLP and social media analysis, with the potential to aid in addressing the issue of racism and promoting diversity and inclusiveness in our society.

REFERENCES

[1]. K. R. Kaiser, D. M. Kaiser, R. M. Kaiser, and A. M. Rackham, ``Using social media to understand and guide the treatment of racist ideology," Global J. Guid. Counseling Schools, Current Perspect., vol. 8, no. 1, pp. 38-49, Apr. 2018.

A. Perrin and M. Anderson. (2018). Share of U.S. Adults Using Social Media, Including Facebook, is Mostly Unchanged Since 2018. [Online]. Available: https://www.pewresearch.org/fact-tank/2019/04/10/share-ofu-s-adults-using-social-media-including-facebook-is-mostly-unchangedsince-2018/

[2]. M. Ahlgren. 40C Twitter Statistics & Facts. Accessed: Sep. 1, 2021. [Online]. Available: https://www.websitehostingrating.com/twitterstatistics/

[3]. D. Arigo, S. Pagoto, L. Carter-Harris, S. E. Lillie, and C. Nebeker, ``Using social media for health research: Methodological and ethical considerations for recruitment and intervention delivery," Digit. Health, vol. 4, Jan. 2018, Art. no. 205520761877175.

[4]. A.-M. Bliuc, N. Faulkner, A. Jakubowicz, and C. McGarty, ``Online networks of racial hate: A systematic review of 10 years of research on cyberracism," Comput. Hum. Behav., vol. 87, pp. 75-86, Oct. 2018.

[5]. M. A. Price, J. R. Weisz, S. McKetta, N. L. Hollinsaid, M. R. Lattanner, A. E. Reid, and M. L. Hatzenbuehler, ``Meta-analysis: Are psychotherapies less effective for black youth in communities with higher levels of anti-black racism?" J. Amer. Acad. Child Adolescent Psychiatry, 2021, doi: 10.1016/j.jaac.2021.07.808.[Online].
Available:https://www.sciencedirect.com/science/article/pii/S0890856721012818

[6]. D. Williams and L. Cooper, ``Reducing racial inequities in health: Using what we already know to take action," Int. J. Environ. Res. Public Health, vol. 16, no. 4, p. 606, Feb. 2019.

[7]. Y. Paradies, J. Ben, N. Denson, A. Elias, N. Priest, A. Pieterse, A. Gupta, M. Kelaher, and G. Gee, ``Racism as a determinant of health: A systematic review and meta-analysis," PLoS ONE, vol. 10, no. 9, Sep. 2015, Art. no. e0138511.

[8]. J. C. Phelan and B. G. Link, ``Is racism a fundamental cause of inequalities in health?" Annu. Rev. Sociol., vol. 41, no. 1, pp. 311-330, Aug. 2015.

[9]. D. R. Williams, ``Race and health: Basic questions, emerging directions," Ann. Epidemiol., vol. 7, no. 5, pp. 322-333, Jul. 1997.

[10]. Z. D. Bailey, N. Krieger, M. Agénor, J. Graves, N. Linos, and M. T. Bassett, ``Structural racism and health inequities in the USA: Evidence and interventions," Lancet, vol. 389, no. 10077, pp. 1453-1463, Apr. 2017.

[11]. D. R. Williams, J. A. Lawrence, B. A. Davis, and C. Vu, ``Understanding how discrimination can affect health," Health Services Res., vol. 54, no. S2, pp. 1374-1388, Dec. 2019.

[12]. C. P. Jones, ``Levels of racism: A theoretic framework and a gardener's tale," Amer. J. Public Health, vol. 90, no. 8, p. 1212, 2000.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 12, April 2025
Impact Factor: 7.67

[13]. S. Forrester, D. Jacobs, R. Zmora, P. Schreiner, V. Roger, and C. I. Kiefe, ``Racial differences in weathering and its associations with psychosocial stress: The CARDIA study," SSM-Population Health, vol. 7, Apr. 2019, Art. no. 100319.

[14]. B. J. Goosby, J. E. Cheadle, and C. Mitchell, ``Stress-related biosocial mechanisms of discrimination and African American health inequities," Annu. Rev. Sociol., vol. 44, no. 1, pp. 319340, Jul. 2018.

[15]. AT. Geronimus, M. Hicken, D. Keene, and J. Bound, ```Weathering' and age patterns of allostatic load scores among blacks and whites in the United States," Amer. J. Public Health, vol. 96, no. 5, pp. 826833, 2006.

[16]. Z. Papacharissi, ``Affective publics and structures of storytelling: Sentiment, events and mediality," Inf., Commun. Soc., vol. 19, no. 3, pp. 307-324, Mar. 2016.

[17]. G. Bouvier, ``How journalists source trending social media feeds: A critical discourse perspective on Twitter," Journalism Stud., vol. 20, no. 2, pp. 212231, Jan. 2019.

[18]. M. KhosraviNik, ``Social media critical discourse studies (SM-CDS)," in The Routledge Handbook of Critical Discourse Studies. London, U.K.: Routledge, 2017, pp. 582-596.

[19]. T. T. Nguyen, S. Criss, A. M. Allen, M. M. Glymour, L. Phan, R. Trevino, S. Dasari, and Q. C. Nguyen, Pride, love, and Twitter rants: Combining machine learning and qualitative techniques to understand what our tweets reveal about race in the US," Int. J. Environ. Res. Public Health, vol. 16, no. 10, p. 1766, May 2019.

[20]. D. T. Goldberg, Are we all Postracial Yet?. Hoboken, NJ, USA: Wiley, 2015.

[21]. Ashraf, S. Hur, and Y. Park, ``Application of deep convolutional neural networks and smartphone sensors for indoor localization," Appl. Sci., vol. 9, no. 11, p. 2337, 2019.

[22]. M. Umer, I. Ashraf, S. Ullah, A. Mehmood, and G. S. Choi, ``COVINet: A convolutional neural network approach for predicting COVID-19 from chest X-ray images," J. Ambient Intell. Humanized Comput.,pp. 1-13, Jan. 2021. [Online]. Available: https://doi.org/10.1007/s12652-021-02917-3

[23]. M. Khalid, I. Ashraf, A. Mehmood, S. Ullah, M. Ahmad, and G. S. Choi, ``GBSVM: Sentiment classication from unstructured reviews using ensemble classifier," Appl. Sci., vol. 10, no. 8, p. 2788, Apr. 2020.

[24]. V. Rupapara, F. Rustam, H. F. Shahzad, A. Mehmood, I. Ashraf, and G. S. Choi, ``Impact of SMOTE on imbalanced text features for toxic comments classification using RVVC model," IEEE Access, vol. 9, pp. 78621-78634, 2021.

[25]. S. Ranjan and S. Sood, ``Social network investor sentiments for predicting stock price trends," Int. J. Sci. Res. Rev., vol. 7, no. 2, pp. 90-97, 2019.

[26]. M. S. Neethu and R. Rajasree, ``Sentiment analysis in Twitter using machine learning techniques," in Proc. 4th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT), Jul. 2013, pp. 1-5.

[27]. Perifanos and D. Goutsos, ``Multimodal hate speech detection in Greek social media," Multimodal Technol. Interact., vol. 5, no. 7, p. 34, 2021.

[28]. Aljarah, M. Habib, N. Hijazi, H. Faris, R. Qaddoura, B. Hammo, M. Abushariah, and M. Alfawareh, ``Intelligent detection of hate speech in arabic social network: A machine learning approach," J. Inf. Sci., vol. 47, no. 3, May 2020, Art. no. 0165551520917651.

[29]. S. Goswami, M. Hudnurkar, and S. Ambekar, ``Fake news and hate speech detection with machine learning and NLP," PalArch's J. Archaeol. Egypt/Egyptol., vol. 17, no. 6, pp. 4309-4322, 2020.

[30]. R. Alshalan and H. Al-Khalifa, ``A deep learning approach for automatic hate speech detection in the Saudi Twitter sphere," Appl. Sci., vol. 10, no. 23, p. 8614, Dec. 2020.

[31]. Ketsbaia, B. Issac, and X. Chen, ``Detection of hate tweets using machine learning and deep learning," in Proc. IEEE 19th Int. Conf. Trust, Secur. Privacy Comput. Commun. (TrustCom), Dec. 2020, pp. 751-758.

[32]. T. Putri, S. Sriadhi, R. Sari, R. Rahmadani, and H. Hutahaean, ``A comparison of classification algorithms for hate speech detection," IOP Conf. Ser., Mater. Sci. Eng., vol. 830, Apr. 2020, Art. no. 032006.

[33]. U. Bhandary, ``Detection of hate speech in videos using machine learning," M.S. thesis, Dept. Comput. Sci., San Jose State Univ., San Jose, CA, USA, 2019.

[34]. B. Vidgen and T. Yasseri, ``Detecting weak and strong Islamophobic hate speech on social media," J. Inf. Technol. Politics, vol. 17, no. 1, pp. 66-78, Jan. 2020