DETECTION OF CYBER BULLYING

WAKCHAURE VAISHNAVI NITESH[1], PATIL APURVA AVINASH[2], SHIRSATH GAURAV DILIP[3], BAGUL SARTHAK JAGATRAO[4], DR RAIS ABDUL HAMID KHAN[5]

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING,
SANDIP UNIVERSITY

Abstract

Aspect Based Sentiment Analysis means to recognize perspectives and feeling polarities towards a given viewpoint in audits. Contrasted and general opinion investigation, ABSA can give more point by point and complete data. As of late, ABSA has turn into a significant errand for normal language understanding and has drawn in extensive consideration from both scholarly and industry fields. The opinion extremity of a sentence isn't just settled by its substance yet in addition has a moderately critical connection with the designated angle. Hence, we propose a model for angle based opinion examination which is a blend of CNN and Gated Recurrent Unit, using the neighborhood highlights produced by CNN and the drawn out reliance learned by GRU. Broad investigations have been directed on datasets of inns and vehicles, and results show that the proposed model accomplishes great execution as far as viewpoint extraction and feeling order. Tests additionally show the incredible space extension ability of the mode.

Keywords: Hate Speech, Machine Learning, Encryption, Detection, SVM

Published in IJIRMPS (E-ISSN: 2349-7300), Volume 11, Issue 6, Nov. – Dec. 2023

INTRODUCTION

Aspect-Based Sentiment Analysis (ABSA) is a type of text analysis that categorizes opinions by aspect and identifies the sentiment related to each aspect. The goal here for the ABSA system is to identify the two aspects – design and price – with their related sentiment. In other words, design: positive, price: negative. Aspect-Based Sentiment Analysis (ABSA) is a type of text analysis that categorizes opinions by aspect and identifies the sentiment related to each aspect. By aspects, we consider attributes or components of an entity (a product or a service, in our case). The sentiment polarity of a sentence is not only decided by its content but also has a relatively significant correlation with the targeted aspect. For this reason, we propose a model for aspect-based sentiment analysis which is a combination of Convolutional Neural Network (CNN) and Gated Recurrent Unit (GRU), utilizing the local features generated by CNN and the long-term dependency learned by GRU

LITURATURE SURVEY

"Hate Speech Dataset from Turkish Tweets", ˙Islam Mayda; Yunus Emre Demir etal.,This paper explained that, Today, while the content produced by users on online platforms increases rapidly due to the spread of the internet, hate speech expressions on these platforms also increase similarly. Social media platforms with millions ofusers are especially among the areas where hate speech expressions are shared frequently. Popular social media companies form their own policies within the scope of combating hate speech. However, the size of the data on the internet makes it almost impossible to do this manually. Consequently, especially in recent years, many studies have been conducted on the automatic detection of hate speech. While most of the studies in the literature are on English, there are published studies on hate speech detection in many languages such as German, French, Arabic, Indonesian, Portuguese. One of the main reasons for fewer studies in languages other than English is the smaller number and size of publicly shared hate speech datasets in those

languages. There is a similar situation for Turkish. Therefore, within the scope of the study, a hate speech dataset comprising 10,224 Turkish tweets was generated and shared publicly. Tweets 4 Machine Learning Approach for Hate Speech and Offensive Language Detection for Social Media were labeled as hate, offensive, and none, and tweets tagged as hate were assigned subclass labels such as ethnic, religious, sexist, and political, which express the type of hate. In the first step of the labeling process, two annotators labeled all tweets separately. In the comparison made after this process, it was seen that the agreement rate in the given labels was 92.5. Afterwards, the two annotators discussed the tweets they gave different labels by exchanging ideas and increased the agreement rate to 98.4. For the remaining tweets, the opinion of the third evaluator was sought. After the labeling process, it was seen that the rate of hate speech in the data set was 22.8. This publicly available data set, which is a first for Turkish in terms of its scope and size, is expected to be an important resource for automatic hate.[1]

”Multi-modal Hate Speech Detection using Machine Learning”, Fariha Tahosin Boishakhi; Ponkoj Chandra Shill etal., This paper presents, With the continuous growth of internet users and media content, it is very hard to track down hateful speech in audio and video. Converting video or audio into text does not detect hate speech accurately as human sometimes uses hateful words as humorous or pleasant in sense and also uses different voice tones or show different action in the video. The state-of-the-art hate speech detection models were mostly developed on a single modality. In this research, a combined approach of multi-modal system has been proposed to detect hate speech from video contents by extracting feature images, feature values extracted from the audio, text and used machine learning and Natural language processing..[2]

”Usage of user hate speech index for improving hate speech detection in Twitter posts” , Ehlimana Krupalija etal., this paper studied that ,Social media is an important source of real-world data for sentiment analysis. Hate speech detection models can be trained on data from Twitter and then utilized for content filtering and removal of posts which contain hate speech. This work proposes a new algorithm for calculating user hate speech index based on user post history. Three available datasets were merged for the purpose of acquiring Twitter posts which contained hate speech. Text preprocessing and tokenization was performed, as well as outlier removal and class balancing. The proposed algorithm was used for determining hate speech index of users who posted tweets from the dataset. The preprocessed dataset was used for training and testing multiple machine learning models: k-means clustering without and with principal component analysis, na¨ıve Bayes, decision tree and random forest. Four different feature subsets of the dataset were used for model training and testing. Anomaly detection, data transformation and parameter tuning were used in an attempt MET’s Institute of Engineering 5 Machine Learning Approach for Hate Speech and Offensive Language Detection for Social Media to improve classification accuracy. The highest F1 measure was achieved by training the model using a combination of user hate speech index and other user features. The results show that the usage of user hate speech index, with or without other user features, improves the accuracy of hate speech detection..[3]

”A Detection of hate speech in Social media memes: A comparative Analysis”, Ajay Nayak, etal., This paper presents, This work projects light upon the challenges of hate speech detection in memes and demonstrates the various machine learning model to automatically detect hate in the internet memes. Memes are the visual content shared on the social media in the form of combination of picture and some textual phrases to depict light humour or jokes. However, some images in the form of memes can also be used to convey misinformation and hate, so their early automatic detection is necessary to stop the hate spreading to wide range of users or population which may cause unrest and harm to human life and property. In this paper, the hateful meme dataset by Facebook AI has been used to test the various unimodal and a multimodal approach to baseline performance for these models and highlight the challenges these hate memes pose to the community..[4]

”Political Hate Speech Detection and Lexicon Building: A Study in Taiwan”, Chih-Chien Wang, etal., In this paper, here is the minimal restriction to users’ speech in cyberspace. The Internet provides a space where people can freely present their speech, which puts a Utopian sense of freedom of speech into practice. However, the appearance of hate speech is a significant side effect of online freedom of speech. Some users use hate speech to attack others, making the attacked targets uncomfortable. The proliferation of hate speech poses severe challenges to cyber society. Users may hope that social media platforms and online communities promote anti-hate speech. However, hate speech detection is still a developing technology that requires system developers to create a method to detect unacceptable hate speech while maintaining the online

freedom of speech environment. No excellence detection approach has yet been proposed, although some literature has focused on it. The current study proposes an approach to build a political hate speech lexicon and train artificial intelligence classifiers to detect hate speech. Our academic and practical contributions include the collection of a Chinese hate speech dataset, creating a Chinese hate speech lexicon, and developing both a deep learning-based and a lexicon-based approach to detect Chinese hate speech. Although we focus on Chinese hate speech detection, our proposed.

AIM & OBJECTIVES

• To Train the Model based on Reviews available in Dataset for Sentiment Prediction.
• To develop a system which is more Security.
• To develop a system which take less time.
• User will detect Negative material from anywhere
• System provide security to user data.
• We are providing easy to used system as compare to existing system. • Time Complexity reduce

MOTIVATION

To detect hate speech messages or comments in social media sites, with machine learning algorithms such as natural language processing (NLP) & Sentiment Analysis (SA)

SYSTEM ARCHITECTURE

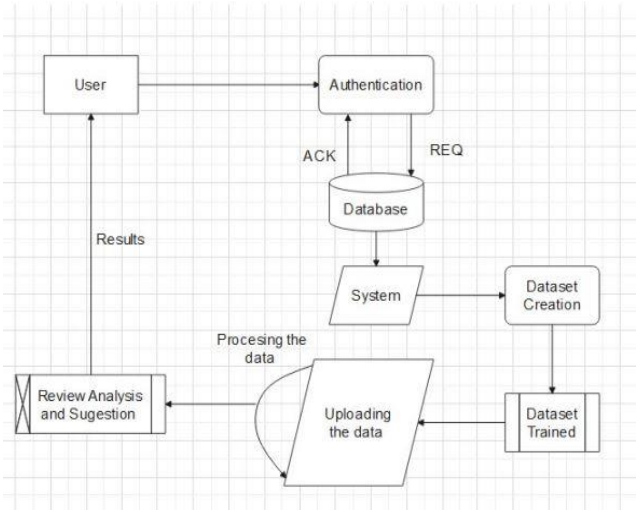

Fig -1: System Architecture Diagram

APPLICATION:

• In defense
• In some rural areas
• Research

FUNCTIONAL & NON-FUNCTIONAL REQUIREMENTS

Functional requirements: may involve calculations, technical details, data manipulation and processing and other specific functionality that define what a system is supposed to accomplish. Behavioral requirements describe all the cases where the system uses the functional requirements; these are captured in use cases.

Nonfunctional Requirements: (NFRs) define system attributes such as security, reliability, performance, maintainability, scalability, and usability. They serve as constraints or restrictions on the design of the system across the different backlogs.

Functional requirements

- Registration
- User Login
- Creation of database: Users Mandatory Information

Design Constraints:

1. Database

2. Operating System
3. Web-Based Non-functional Requirements

Security:

1. User Identification
2. Login ID
3. Modification

Performance Requirement:

1. Response Time
2. Capacity
3. User Interface
4. Maintainability
5. Availability

SYSTEM REQUIREMENTS

Hardware Requirements

• AMD/Intel i3 Processor or above Processor
• 4GB RAM for application development
• 80 GB or above Hard Disk

Software Requirements

• Windows 7 or above
• Vscode, Xamp
• Python

CONCLUSION

Hence we are overcoming the drawback of existing system, we are providing the better solution than existing system in affordable cost. We proposed a system which is used to identify the aspect sentiments detection using CNN algorithm, which is based deep learning. Aspect Sentiment Analysis is often performed on aspect to help businesses monitor brand and product sentiment in customer feedback, and understand customer needs

REFERENCES

[1] NARISA ZHAO , HUAN GAO , XIN WEN , AND HUI LI, Received January 12, 2021, accepted January 15, 2021, date of publication January 19, 2021, date of current version January 27, 2021. ”Combination of Convolutional Neural Network and Gated Recurrent Unit for Aspect-Based Sentiment Analysis”.

[2] Xingyou Wang1, Weijie Jiang2, Zhiyong Luo, Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, pages 2428–2437, Osaka, Japan, December 11-17 2016. ”Combination of Convolutional and Recurrent Neural Network for Sentiment Analysis of Short Texts”.

[3] MohammadAL-Smadi, Mahmoud M.Hammad, Sa’ad A.Al-Zboon, SajaALTawalbeh, ErikCambria, “ Gated Recurrent Unit with Multilingual Universal Sentence Encoder for Arabic Aspect-Based Sentiment Analysis” Knowledge based Systems is an international and interdisciplinary journal in the field of artificial intelligence, 2021.

[4] Mohsen Ghorbani, Mahdi Bahaghighat , Qin Xin and Figen Ozen, Journal of ¨ Cloud Computing: Advances, Systems and Applications (2020), ConvLSTMConv network: a deep learning approach for sentiment analysis in cloud computing.

[5] E. Walter, Cambridge advanced Learner’s dictionary, Cambridge Univ. Press, 2012 .

[6] W. Warner and J. Hirschberg, ”Detecting Hate Speech on the World Wide Web” in , New York, NY:Columbia University, Department of Computer Science, pp. 10027D.

[7] S. Macavaney, Hao-Ren Yao-E. Yang, K. Russell, N. Goharian and O. Frieder, ”PLOS ONE “Hate speech detection: Challenges and solutions””, 2019 Master Business Analytics Department of Mathematics Faculty of Science, August, 2018.