IJETIR

ISSN (Online) 2583-0554

International Journal of Emerging Technologies and Innovative Research (IJETIR)

Volume 4, Issue 6, June 2024

Impact Factor: 5.731

Weakly Supervised Deep Embedding for Product Review Sentiment Analysis

Sandeep P[1] and Dr. D. R. Krithika[2]

II MCA Student, Department of Computer Applications[1]

Assistant Professor, Department of Computer Applications[2]

22304138@vistas.ac.in and krithika.scs@velsuniv.ac.in

Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai, Tamil Nadu, India

Abstract: *Online reviews have become an important source of information for users before making an informed purchase decision. Early reviews of a product tend to have a high impact on the subsequent product sales. In this project, we take the initiative to study the behavior characteristics of early reviewers through their posted reviews on two real-world large e-commerce platforms, i.e., Amazon and Yelp. In specific, we divide product lifetime into three consecutive stages, namely early, majority and laggards. A user who has posted a review in the early stage is considered as an early reviewer. We quantitatively characterize early reviewers based on their rating behaviors, the helpfulness scores received from others and the correlation of their reviews with product popularity. We have found that (1) an early reviewer tends to assign a higher average rating score; and (2) an early reviewer tends to post more helpful reviews. Our analysis of product reviews also indicates that early reviewers' ratings and their received helpfulness scores are likely to influence product popularity. By viewing review posting process as a multiplayer competition game, we propose a novel margin-based embedding model for early reviewer prediction. Extensive experiments on two different e-commerce datasets have shown that our proposed approach outperforms a number of competitive baselines. In our project we have used algorithms like Decision Tree (DT) and Multi Layer Perceptron (MLP). All are measured in terms of accuracy*

Keywords: Decision Tree (DT), Multi Layer Perceptron (MLP)

I. INTRODUCTION

MACHINE LEARNING

In the statistical context, Machine Learning is defined as an application of artificial intelligence where available information is used through algorithms to process or assist the processing of statistical data. While Machine Learning involves concepts of automation, it requires human guidance. Machine Learning involves a high level of generalization in order to get a system that performs well on yet unseen data instances.Machine learning is a relatively new discipline within Computer Science that provides a collection of data analysis techniques. Some of these techniques are based on well established statistical methods (e.g. logistic regression and principal component analysis) while many others are not.Web enable users to keep in touch with friends, relatives, family members, and people with similar interests, profession, and objectives. In addition, they allow users to interact with one another and form communities.A user can become a member of an Twitter by registering and providing details, such as name, birthday, gender, and other contact information. Although a large number of web exist on the web, Facebook and Twitter are among the most popular webs are included in the list of the top 10 websites[1] around the worldwide. Twitter allows the users to follow their favorite politicians, athletes, celebrities, and news channels, and to subscribe to their content without any hindrance. Through following activity, a follower can receive status updates of subscribed account. Although Twitter and other OSNs are mainly used for various benign purposes, their open nature, huge user base, and real-time message proliferation have made them lucrative targets for cyber criminals and social bots.

IJETIR
ISSN (Online) 2583-0554
International Journal of Emerging Technologies and Innovative Research (IJETIR)
Volume 4, Issue 6, June 2024
Impact Factor: 5.731

II. LITERATURE SURVEY

Trending Topics based Spam Detection from the Social Media Aditi Gupta ; NeetikaRaina ; BhartiJha ; Vinay Kumar Jain IEEE 2021.

With an increasing usage of social media to exchange, share and store information, cybercriminals also get attracted to it, to take advantage of the network for their illegal and unethical benefits. Fake online accounts pop up every day. Spammers are the users behind the screen who share unsolicited and irrelevant texts to a huge number of users with an intent of advertising some product or to make people to click on unsecured links and infecting user's system usually to make money (click bait). They often use Trending topics on social media as a medium to spam. Sometimes, spam and fake trending is created by Spammers and many a times spammers use Trending topics to lure victims into clicking them [1].

A Neural Network-Based Ensemble Approach for Spam Detection in Twitter SreekanthMadisetty ; Maunendra Sankar Desarkar IEEE 2021.

As the social networking sites get more popular, spammers target these sites to spread spam posts. Twitter is one of the most popular online social networking sites where users communicate and interact on various topics. Most of the current spam filtering methods in Twitter focus on detecting the spammers and blocking them. However, spammers can create a new account and start posting new spam tweets again. So there is a need for robust spam detection techniques to detect the spam at tweet level. These types of techniques can prevent the spam in real time. To detect the spam at tweet level, often features are defined, and appropriate machine learning algorithms are applied in the literature [2].

Spam Reviews Detection Framework Based on Heterogeneous Information Network (HIN) Arpita Kunne ; R. Roopalakshmi ; Suviksha. V. Shetty ; S.V. Ananya ; Pankaj Devidas Divgi ; Manasa IEEE 2021.

Nowadays social media plays an important role in our day-to-day activities. Specifically, in the past few years, online social websites such as Facebook, Twitter and WhatsApp are evolving as one of the major sources of communication for internet users, in order to keep in touch with their friends. However, Spam reviews generated on a website results in huge financial gain only for competitors, whereas it is a major loss for both customers and organization. In the literature, the existing techniques for Spam reviews detection suffer due to issues such as limited datasets and lack of proper classification methods, which results in inefficiency of the systems. In order to solve these problems, this paper proposes a new framework, which models the given review dataset, using Heterogeneous Information Network (HIN) concept and solves the spam detection problem by means of clearly identifying the spam reviews present in a website [3].

Online Spammer Detection using User-Neighbor Relationship SihyunJeong ; Chong-kwon Kim IEEE 2021.

Today, online social media is being used as a primary means of interaction among people rather than any other communication medium. People use social media for socializing and collecting and sharing information, which presumes online social media is always reliable. However, there are a lot of advertisements or cyberattacks that take advantage of the impact of online social media. For example, spam, phishing, and advertising through influence manipulation. They are not only directly damaging users, but they are also giving indirect damages that will cause damage to the credibility of online social media to decline [4].

A Novel Approach for Youtube Video Spam Detection using Markov Decision Process SimranKanodia ; RachnaSasheendran ; VinodPathari IEEE 2021.

Social networking websites have become an integral part of the day to day lives of people. People turn to social media for interacting with other people, sharing ideas, gaining knowledge, for entertainment and staying informed about the events happening in the rest of the world. Among these sites, YouTube has emerged as the most popular website for sharing and viewing video content. This popularity of YouTube has also attracted spammers, who upload videos with the sole purpose of polluting the system content and causing dissatisfaction among other viewers. These spam videos may be unrelated to their title or may contain pornographic content. Therefore, it is very important to find a way to detect these videos and report them before they are viewed by innocent users [5].

IJETIR
ISSN (Online) 2583-0554
International Journal of Emerging Technologies and Innovative Research (IJETIR)
Volume 4, Issue 6, June 2024
Impact Factor: 5.731

III. RESULTS AND DISCUSSION

1. PROBLEM STATEMENT

- No user connection at various levels
- Privacy violation.
- Unwanted panic creation.
- Unwanted news spread among people wasting time.

2. PROPOSED SYSTEM

Multilayer Peceptron (MLP): A **multilayer perceptron** (**MLP**) is a class of feed forward artificial neural network (ANN). The term MLP is used ambiguously, sometimes loosely to mean *any* feed forward ANN, sometimes strictly to refer to networks composed of multiple layers of perceptrons (with threshold activation); see § Terminology. Multilayer perceptrons are sometimes colloquially referred to as "vanilla" neural networks, especially when they have a single hidden layer. An MLP consists of at least three layers of nodes: an input layer, a hidden layer and an output layer. Except for the input nodes, each node is a neuron that uses a nonlinear activation function. MLP utilizes a supervised learning technique called back propagation for training. Its multiple layers and non-linear activation distinguish MLP from a linear perceptron. It can distinguish data that is not linearly separable.

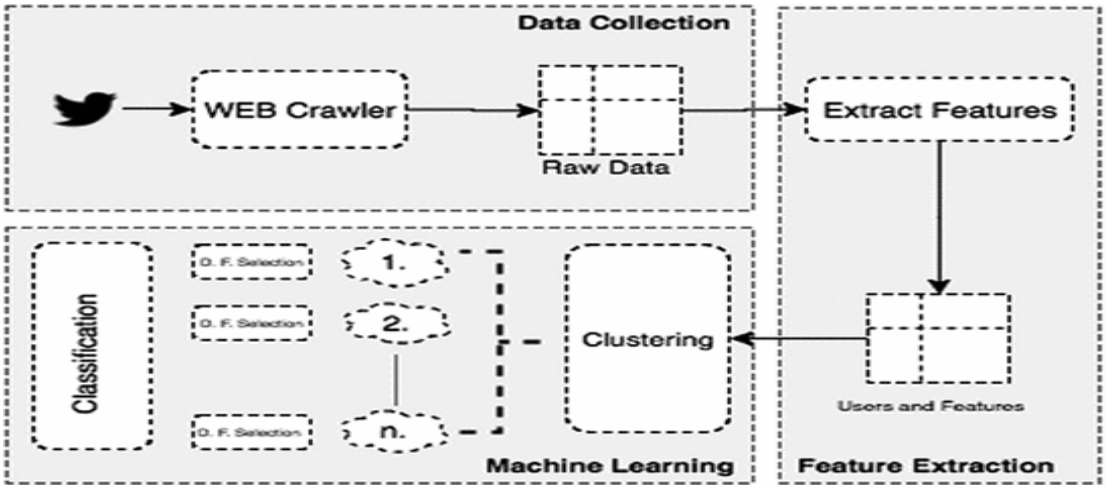

Fig 1. ARCHITECTURE DIAGRAM

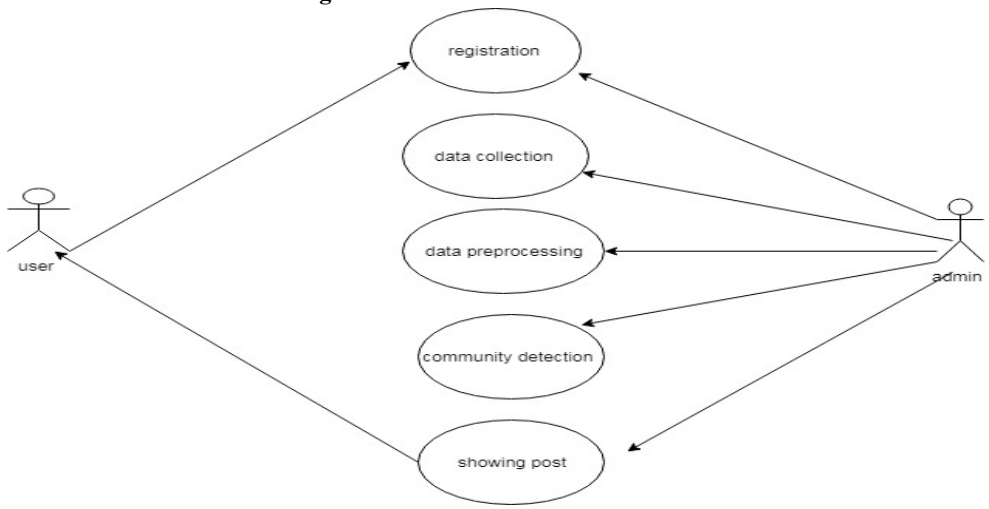

Fig 2. USECASE DIAGRAM

IV. CONCLUSION

Social media networks are widely used communication channels to exchange information all over the world. Along with the benefits of social media networks, some of spammers spread unwanted information into network. This data misguide the genuine users.The experiments conducted on the Twitter dataset demonstrate the effectiveness of the proposed model. The prediction ability of the proposed model is further verified on the opinion word prediction task. Based on the learned influence, we explore the expression styles of users with different influence powers, which provide the valuable information for many users.

REFERENCES

[1]. C. Grier, K. Thomas, V. Paxson, and M. Zhang, "@spam: The underground on 140 characters or less," in proc. ACM conf. Computer communication security, 2019, pp. 27-37.

[2]. Y Boshmaf, I.Muslukhov, K Beznoson, and M. Ripeanu, " design and analysis of social botnet," computer networks, vol. 57, no. 2, pp. 556- 578, 2019.

[3]. Arora, Harsha, Govinda Murali Upadhyay, "A framework for the detection of Suspicious Discussion on online forum using integrated approach of support vector machine and particle Swarm Optimization", international Journal of Advanced research in computer science, 2019.

[4]. M. Kirby, L. Sirovich, "Application of the Karhunen loeve Procedure for the Characterization of Human Faces", IEEE Transactions on Pattern analysis and Machine Intelligence, vol. 12, no. 1. Pp. 103-107, 2019.

[5]. M. McCord, M. Chuah, "Spam Detection on Twitter Using Traditional Classifiers", pp. 175-186, Springer, 2019.