SPAMMER DETECTION AND FAKE USER IDENTIFICATION ON SOCIAL NETWORKS

A PROJECT REPORT

Submitted by

SRIHARIHARAN.T (1714172)

SHANMUGASURYA.S (1714169)

in partial fulfillment of the requirement

for the award of the degree

o

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

K.S. RANGASAMY COLLEGE OF TECHNOLOGY
(An Autonomous Institution, affiliated to Anna University Chennai and Approved by AICTE, New Delhi)

TIRUCHENGODE – 637 215

MAY 2021

K.S. RANGASAMY COLLEGE OF TECHNOLOGY
TIRUCHENGODE - 637 215

BONAFIDE CERTIFICATE

Certified that this project report titled **“SPAMMER DETECTION AND FAKE USER IDENTIFICATION ON SOCIAL NETWORKS”** is the bonafide work of **SRIHARIHARAN.T (1714172), SHANMUGASURYA.S (1714169)** who carried out the project under my supervision. Certified further, that to the best of my knowledge the work reported herein does not form part of any other project report or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

DR.S.MADHAVI M.E., PhD.	**DR.P.SENTHIL RAJA M.E , PhD.**
HEAD OF THE DEPARTMENT	**SUPERVISOR**
PROFESSOR	PROFESSOR
Department of computer science and engineering	Department of computer science and engineering
K.S. Rangasamy College of Technology	K.S. Rangasamy College of Technology
Tiruchengode - 637 215	Tiruchengode - 637 215

Submitted for the viva-voce examination held on ………………

Internal Examiner **External Examiner**

DECLARATION

I/We jointly declare that the project report on **“SPAMMER DETECTION AND FAKE USER IDENTIFICATION ON SOCIAL NETWORKS”** is the result of original work done by me/us and best of my/our knowledge, similar work has not been submitted to **“ANNA UNIVERSITY CHENNAI”** for the requirement of Degree of Bachelor of ngineering. This project report is submitted on the partial fulfilment of the requirement of the award of Degree of Bachelor of Engineeering.

Signature

T. Sri

SRIHARIHARAN T

S. Shanmuga Surya.

SHANMUGASURYA S

Date:

Place: Tiruchengode

ACKNOWLEDGEMENT

We wish to express our sincere gratitude to our honourable Chairman **LION Dr. K.S.RANGASAMY M.J.F.,** and Vice Chairman **R.SRINIVASAN B.B.M., MISTE.,** for providing immense facilities at our institution.

We would like to express special thanks of gratitude to our Chief Executive Officer of **Dr.K.THYAGARAJAH M.E., PhD., SMIEE, MISTE,** who has been the key spring of motivation to us throughout the completion of our course and project work.

We are very proudly rendering our thanks to our Principal **Dr.R.GOPALAKRISHNAN M.E., PhD.,** for the facilities and the encouragement given by him to the progress and completion of our project.

We proudly render our immense gratitude to Director of Computer Science and Engineering **Dr.B.G.GEETHA M.E., PhD.,** for her effective leadership, encouragement and guidance in the project.

We proudly render our immense gratitude to the Head of the Computer Science and Engineering **Dr.S.MADHAVI M.E., PhD.,** for her effective leadership, encouragement and guidance in the project.

We are highly indebted to provide our heart full thanks to our supervisor **Dr.P.SENTHIL RAJA ME., PhD** Associate Professor for his valuable ideas, encouragement and supportive guidance throughout the project.

We are highly indebted to provide our heart full thanks to our Project coordinator **Dr.R.GOPINATH M.E., PhD.,** Associate Professor for his valuable ideas, encouragement and supportive guidance throughout the project.

We wish to extend our sincere thanks to all faculty members of our Department of Computer Science and Engineering for their valuable suggestions, kind co-operation and constant encouragement for successful completion of this project.

We wish to acknowledge the help received from various Departments and various individuals during the preparation and editing stages of the manuscript.

.

ABSTRACT

As e-commerce is growing and becoming popular day-by-day, the number of reviews received from customer about any product grows rapidly. People nowadays heavily rely on reviews before buying anything. Product reviews play an important role in deciding the sale of a particular product on the ecommerce websites or applications like Flipkart, Amazon, Snapdeal, etc. In this paper, we propose a framework to detect fake product reviews or spam reviews by using Opinion Mining . The Opinion mining is also known as Sentiment Analysis. In sentiment analysis, we try to figure out the opinion of a customer through a piece of text.

The proposed method called VWNB-FIUT (Value Weighted Naïve Bayes with Frequent Pattern Ultra Metric Tree) automatically classifies users' reviews into "suspicious", "clear" and "hazy" categories by phase-wise processing. The hazy category recursively eliminates elements into suspicious or clear. This results into richer detection and be useful to business organization as well as to customers. Business organization can monitor their product selling by analyzing and understanding what the customers are saying about products.

This can help customers to purchase valuable product and spend their money on quality products. Finally end users see that each individual review with polarity scores and credibility score annotated on it. We first take the review and check if the review is related to the specific product with the help of VWNB. We use Spam dictionary to identify the spam words in the reviews by using FIUT. In Text Mining we apply several algorithms and on the basis of these algorithms we get the specific results.

.

LIST OF SYMBOLS AND ABBREVIATIONS

SYMBOLS

G	-	Acceleration due to gravity	m/s^2
A	-	Area	m^2

GREEK SYMBOLS

P	-	Density of air	kg/m^3
η	-	Efficiency	%

ABBREVIATIONS

VWNB-FIUT	-	Value Weighted Naïve Bayes with Frequent Pattern Ultra Metric Tree
LU	-	Labelled and unlabeled
EM	-	expectation–maximization
MAP	-	maximum posterior
ASM	-	Author Spam city Model
HPSD	-	hybrid PU-learning-based Spammer Detection
CQA	-	Community Question Answering
BOW	-	Bag of Words

TABLES

CHAPTER 1

INTRODUCTION

1.1 ECOMMRCE PLATFORM

An ecommerce platform is a software application that allows online businesses to manage their website, marketing, sales, and operations. E-commerce is the activity of buying or selling of products on online services or over the Internet. Electronic commerce draws on technologies such as mobile commerce, electronic funds transfer, supply chain management, Internet marketing, online transaction processing, electronic data interchange (EDI), inventory management systems, and automated data collection systems.

1.2 OPINION MINING

Product inference mining is a process of tracking the mood of the public about a particular product . Opinions can be essential when it's use to make a decision or choose among multiple option. Information-gathering behavior has always been to find out what other people think. The availability of opinion-rich resources such as online review sites and personal blogs, and challenges arise, to understand the opinions of others people.

1.2.1 Levels of product inference Mining

Product inference mining is extracting people's opinion from the web. It is also known as sentiment analysis. There are three tasks for opinion mining [14]

- Document-level opinion mining
- Sentence-level opinion mining
- Phrase-level opinion mining

1.2.2 TASKS IN OPINION MINING

The area of opinion analysis is to predict the polarity of a piece of opinion text as positive or negative. Product inference analysis tasks unnoticed due to lack of popularity. Here, the tasks related to opinion analysis are,[4]

- Subjectivity Detection
- Sentiment Prediction
- Aspect Based Sentiment Summarization
- Contrastive Viewpoint Summarization
- Text Summarization for Opinions
- Predicting Helpfulness of Online Comments/Reviews
- product inference-Based Entity Ranking

1.2.2.1 Subjectivity Detection

The task is about determining a piece of text actually contains opinions or not. It is not much about determining the polarity of the text.

1.2.2.2 Sentiment Prediction

Sentiment task is about predicting the polarity of a piece of text usually positive or negative. People have studied sentiment prediction at the document level, sentence level and phrase level. This is an extremely popular task in the field of product inference Analysis.

1.2.2.3 Aspect Based Sentiment Summarization

This task goes beyond sentiment prediction the goal is to provide a summary in the form of star ratings or scores on each of these features. So the task involves finding features and then discovering the sentiments for each feature.

1.2.2.4 Contrastive Viewpoint Summarization

This task is about try to highlight contradiction in opinions were present. In contrastive viewpoints highlighted, people can get a better understanding of the opinions and under which condition it holds.

1.2.2.5 Text Summarization for Opinions

Instead of generating structured summaries of opinions, another useful summary format is to generate textual summaries. For example, a few sentences summarize the reviews of a product or a set of phrases acting as summaries.

1.3 SPAMMER GROUP

We focus on group spam, which has not been studied so far. A spammer group refers to a group of reviewers who works together writing fake reviews to promote or demote a set of target products. Spammer groups are very damaging due to their sheer sizes. It is well-known that many online reviews are not written by genuine users of products, but by spammers who write fake reviews to promote or demote some target products. Although some existing works have been done to detect fake reviews and individual spammers, to our knowledge, no work has been done on detecting spammer groups. This work focuses on this task and proposes an effective technique to detect such groups.

1.4 SUPERVISED LEARNING

In supervised learning, the learning algorithm uses labelled training examples from every class to generate a classification function. One of the drawbacks of this classic paradigm is that a large number of labelled examples are needed in order to learn accurately. Since labelling is often done manually, it can be very labour intensive and time consuming. In this chapter, we study two partially supervised learning tasks. As their names suggest, these two learning tasks do not need full supervision, and thus are able to reduce the labelling effort.

The first is the task of learning from labelled and unlabeled examples, which is commonly known as semi supervised learning. In this chapter, we also call it LU learning (L and U stand for "labelled" and "unlabeled" respectively). In this learning setting, there is a small set of labelled examples of every class, and a large set of unlabeled examples. The objective is to make use of the unlabeled examples to improve learning.

1.5 LEARNING FROM POSITIVE AND UNLABELED DATA

Learning from positive and unlabeled data or PU learning is the setting where a learner only has access to positive examples and unlabeled data. The assumption is that the unlabeled data can contain both positive and negative examples.

1.6 INFORMATION EXTRACTION

A Reliability Study for Evaluating Information Extraction from Radiology Reports Setting: Twenty-four physician raters from two sites and two specialties. Are in fact incorrect, and a small proportion of the negative answers are incorrect.

1.7 NAÏVE BAYESIAN

Naive Bayesian. The Naive Bayesian classifier is based on Bayes' theorem with the independence assumptions between predictors. A Naive Bayesian model is easy to build, with no complicated iterative parameter estimation which makes it particularly useful for very large datasets.

1.8 EM ALGORITHM

In statistics, an expectation–maximization (EM) algorithm is an iterative method to find maximum likelihood or maximum posterior (MAP) estimates of parameters in statistical models, where the model depends on unobserved latent variables.

CHAPTER 2

LITERATURE REVIEW

2.1 IMPACT OF ONLINE CONSUMER REVIEWS ON SALES

F. Zhu and X. Zhang et.al ‘‘Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics,” in this work how product and consumer characteristics moderate the influence of online consumer reviews on product sales using data from the video game industry. The findings indicate that online reviews are more influential for less popular games and games whose players have greater Internet experience.

The differential impact of consumer reviews across products in the same product category and suggests that firms’ online marketing strategies should be contingent on product and consumer characteristics [1].

2.2 TEMPORAL DYNAMICS OF OPINION SPAMMING

K. C. Santosh and A. Mukherjee et .al ‘‘on the temporal dynamics of opinion spamming: Case studies on yelp,’’ In this work recently, the problem of opinion spam has been widespread and has attracted a lot of research attention. While the problem has been approached on a variety of dimensions, the temporal dynamics in which opinion spamming operates is unclear.

How does buffered spamming operate for entities that need spamming to retain threshold popularity and reduced spamming for entities making better success? We analyze these questions in the light of time-series analysis on Yelp. This work analyses discover various temporal patterns and their relationships with the rate at which fake reviews are posted. Building on our analyses, we employ vector auto regression to predict the rate of deception across different spamming policies.

Next, we explore the effect of filtered reviews on (long-term and imminent) future rating and popularity prediction of entities. Our results discover novel temporal dynamics of spamming which are intuitive, arguable and also render confidence on Yelp’s filtering. Lastly, we leverage our discovered temporal patterns in deception detection. Experimental results on large-scale reviews show the effectiveness of our approach that significantly improves the existing approaches [2].

2.3 OPINION SPAM AND ANALYSIS

N. Jindal and B. Liu et.al, ‘‘Opinion spam and analysis,’’ in this work past few years, sentiment analysis and opinion mining becomes a popular and important task. These studies all assume that their opinion resources are real and trustful. However, they may encounter the faked opinion or opinion spam problem. We study this issue in the context of our product review mining system. On product review site, people may write faked reviews, called review spam, to promote their products, or defame their competitors’ products. It is important to identify and filter out the review spam. Previous work only focuses on some heuristic rules, such as helpfulness voting, or rating deviation, which limits the performance of this task. We exploit machine learning methods to identify review spam. Toward the end, we manually build a spam collection from our crawled reviews. We first analyze the effect of various features in spam identification. We also observe that the review spammer consistently writes spam.

This provides us another view to identify review spam: we can identify if the author of the review is spammer. Based on this observation, we provide a two view. Semi-supervised method, co-training, to exploit the large amount of unlabeled data. The experiment results show that our proposed method is effective. Our designed machine learning methods achieve significant improvements in comparison to the heuristic baselines [3].

2.4 LEARNING TO IDENTIFY REVIEW SPAM

F. Li, M. Huang, Y. Yang, and X. Zhu et.al, ‘‘Learning to identify review spam,’ in these work online reviews plays a crucial role in today’s electronic commerce. It is desirable for a customer to read reviews of products or stores before making the decision of what or from where to buy. Due to the pervasive spam reviews, customers can be misled to buy low-quality products, while decent stores can be defamed by malicious reviews. We observe that, in reality, a great portion of the reviewers write only one review .These reviews are so enormous in number that they can almost determine a store’s rating and impression.

In existing methods did not examine this larger part of the reviews. To address this problem, we observe that the normal re- viewers’ arrival pattern is stable and uncorrelated to their rating pattern temporally. In contrast, spam attacks are usually busty and either positively or negatively correlated to the rating. Thus, we propose to detect such attacks via unusually correlated temporal patterns. We identify and construct multidimensional time series based on aggregate statistics, in order to depict and mine such correlations.

In this way, the singleton review spam detection problem is mapped to a abnormally correlated pattern detection problem. We propose a hierarchical algorithm to robustly detect the time windows where such attacks are likely to have happened. The algorithm also pinpoints such windows in different time resolutions to facilitate faster human inspection. Experimental results show that the proposed method is effective in detecting singleton review attacks. We discover that singleton review is a significant source of spam reviews and largely affects the ratings of online stores [4].

2.5 REVIEW SPAM DETECTION VIA TEMPORAL PATTERN DISCOVERY

S. Xie, G. Wang, S. Lin, and P. S. Yu et.al ‘‘Review spam detection via temporal pattern discovery,’’ in this work used by individuals and organizations for their decision making. However, due to the reason of profit or fame, people try to game the system by opinion spamming (e.g., writing fake reviews) to promote or to demote some target products. In recent years, fake review detection has attracted significant attention from both the business and research communities.

However, due to the difficulty of human labeling needed for supervised learning and evaluation, the problem remains to be highly challenging. This work proposes a novel angle to the problem by modeling spam city as latent. An unsupervised model, called Author Spam city Model (ASM), is proposed. It works in the Bayesian setting, which facilitates modeling spam city of authors as latent and allows us to exploit various observed behavioral footprints of reviewers. The intuition is that opinion spammers have different behavioral distributions than non-spammers.

This creates a distributional divergence between the latent population distributions of two clusters: spammers and non-spammers. Model inference results in learning the population distributions of the two clusters. Several extensions of ASM are also considered leveraging from different priors. Experiments on a real-life Amazon review dataset demonstrate the effectiveness of the proposed models which significantly outperform the state-of-the-art competitors [5].

2.6 SPOTTING OPINION SPAMMERS USING BEHAVIORAL FOOTPRINTS

A. Mukherjee et al ‘‘Spotting opinion spammers using behavioral footprints,” in these work User-generated online reviews can play a significant role in the success of retail products, hotels, restaurants, etc. However, review systems are often targeted by opinion spammers who seek to distort the perceived quality of a product by creating fraudulent reviews. We propose a fast and effective framework, fraud eagle, for spotting fraudsters and fake reviews in online review datasets. in this method has several advantages: it exploits the network effect among reviewers and products, unlike the vast majority of existing methods that focus on review text or behavioral analysis, it consists of two complementary steps; scoring users and reviews for fraud detection, and grouping for visualization and sense making, it operates in a completely unsupervised fashion requiring no labeled data, while still incorporating side information if available, and it is scalable to large datasets as its run time grows linearly with network size. We demonstrate the effectiveness of our framework on synthetic and real datasets; where fraud eagle successfully reveal fraud-bots in a large online app review database [6].

2.7 OPINION FRAUD DETECTION IN ONLINE REVIEWS BY NETWORK EFFECTS

L. Akoglu, R. Chandy, and C. Faloutsos, et.al ‘‘Opinion fraud detection in online reviews by network effects,’’ big part of people rely on available content in social media in their decisions (e.g. reviews and feedback on a topic or product). The possibility that anybody can leave a review provides a golden opportunity for spammers to write spam reviews about products and services for different interests.

Identifying these spammers and the spam content is a hot topic of research and although a considerable number of studies have been done recently toward this end, but so far the methodologies put forth still barely detect spam reviews, and none of them show the importance of each extracted feature type. In this study, we propose a novel framework, named Net Spam, which utilizes spam features for modeling review datasets as heterogeneous information networks to map spam detection procedure into a classification problem in such networks. Using the importance of spam features help we to obtain better results in terms of different metrics experimented on real-world review datasets from Yelp and Amazon websites.

The results show that Net Spam outperforms the existing methods and among four categories of features; including review-behavioral, user-behavioral, review linguistic, nuser-linguistic, the first type of features performs better than the other categories [7].

2.8 NET SPAM: A NETWORK-BASED SPAM DETECTION FRAMEWORK FOR REVIEWS IN ONLINE SOCIAL MEDIA

S. Shehnepoor, M. Salehi, R. Farahbakhsh, and N. Crespi et .al ‘‘Net Spam: A network-based spam detection framework for reviews in online social media,’’ In this work Driven by profits, spam reviews for product promotion or suppression become increasingly rampant in online shopping platforms.

This work focuses on detecting hidden spam users based on product reviews. In the literature, there have been tremendous studies suggesting diversified methods for spammer detection, but whether these methods can be combined effectively for higher performance remains unclear. Along this line, a hybrid PU-learning-based Spammer Detection (hPSD) model is proposed in this work. On one hand, hPSD can detect multi-type spammers by injecting or recognizing only a small portion of positive samples, which meets particularly real-world application scenarios. More importantly, hPSD can leverage both user features and user relations to build a spammer classifier via a semi-supervised hybrid learning framework.

Experimental results on amazon data sets with shilling injection show that hPSD outperforms several state-of-the-art baseline methods. In particular, hPSD shows great potential in detecting hidden spammers as well as their underlying employers from a real life Amazon data set. These demonstrate the effectiveness and practical value of hPSD for real-life applications [8].

2.9 SPAMMERS DETECTION FROM PRODUCT REVIEWS: A HYBRID MODEL

Z. Wu, Y. Wang, Y. Wang, J. Wu, J. Cao, and L. Zhang et.al ‘‘Spammers detection from product reviews: A hybrid model,’’in this work Today’s e-commerce is highly depended on increasingly growing online customers’ reviews posted in opinion sharing websites.

This fact, unfortunately, has tempted spammers to target opinion sharing web- sites in order to promote and demote products. To date, different types of opinion spam detection methods have been proposed in order to provide reliable resources for customers, manufacturers and re- searchers. However, supervised approaches suffer from imbalance data due to scarcity of spam reviews in datasets, rating deviation based filtering systems are easily cheated by smart spammers, and content based methods are very expensive and majority of them have not been tested on real data hitherto.

The aim of this work is to propose a robust review spam detection system wherein the rating deviation, content based factors and activeness of reviewers are employed efficiently. To overcome the aforementioned drawbacks, all these factors are synthetically investigated in suspicious time intervals captured from time series of reviews by a pattern recognition technique. The proposed method could be a great asset in online spam filtering systems and could be used in data mining and knowledge discovery tasks as a standalone system to purify product review datasets.

These systems can reap benefit from our method in terms of time efficiency and high accuracy. Empirical analyses on real dataset show that the proposed approach is able to successfully detect spam reviews. Comparison with two of the current common methods, indicates that our method is able to achieve higher detection accuracy (F-Score: 0.86) while removing the need for having specific fields of Meta data and reducing heavy computation required for investigate purposes [9].

2.10 DETECTION OF FAKE OPINIONS USING TIME SERIES

A. Heydari, M. Tavakoli, and N. Salim et.al ‘‘Detection of fake opinions using time series, “in this work Online reviews play a crucial role in helping consumers evaluate and compare products and services. This critical importance of reviews also incentivizes fraudsters (or spammers) to write fake or spam reviews to secretly promote or demote some target products and services. Existing approaches to detecting spam reviews and reviewers employed review contents, reviewer behaviors, star rating patterns, and reviewer-product networks for detection. In this research, we further discovered that reviewers’ posting rates (number of reviews written in a period of time) also follow an interesting distribution pattern, which has not been reported before. That is, their posting rates are bimodal.

Multiple spammers also tend to collectively and actively post reviews to the same set of products within a short time frame, which we call co-bursting. Furthermore, we found some other interesting patterns in individual reviewers’ temporal dynamics and their co-bursting behaviors with other reviewers. Inspired by these findings, we first propose a two-mode Labeled Hidden Markov Model to model spamming using only individual reviewers’ review posting times. We then extend it to the Coupled Hidden Markov Model to capture both reviewer posting behaviors and co-bursting signals.

In these experiments show that the proposed model significantly outperforms state-of-the-art baselines in identifying individual spammers. Furthermore, we propose a co bursting network based on co-bursting relations, which helps detect groups of spammers more effectively than existing approaches[10].

2.11 BIMODAL DISTRIBUTION AND CO-BURSTING IN REVIEW SPAM DETECTION

H. Li et al., ''Bimodal distribution and co-bursting in review spam detection,'' in this work As the rapid development of China's e-commerce in recent years and the underlying evolution of adversarial spamming tactics, more sophisticated spamming activities may carry out in Chinese review websites. Empirical analysis, on recently crawled product reviews from a popular Chinese e- commerce website, reveals the failure of many state-of the- art spam indicators on detecting collusive spammers.

Two novel methods are then proposed: a KNN-based method that considers the pairwise similarity of two reviewers based on their group-level relational information and selects *k* most similar reviewers for voting; a more general graph-based classification method that jointly classifies a set of reviewers based on their pairwise transaction correlations. Experimental results show that both our methods promisingly outperform the indicator-only classifiers in various settings [11].

2.12 SPOTTING FAKE REVIEWER GROUPS IN CONSUMER REVIEWS

A. Mukherjee, B. Liu, and N. Glance, ''Spotting fake reviewer groups in consumer reviews,'' in this work Opinionated social media such as product reviews are now widely used by individuals and organizations for their decision making. However, due to the reason of profit or fame, people try to game the system by opinion spamming (e.g., writing fake reviews) to promote or demote some target products. For reviews to reflect genuine user experiences and opinions, such spam reviews should be detected. Prior works on opinion spam focused on detecting fake reviews and individual fake reviewers. However, a fake reviewer group (a group of reviewers who work collaboratively to write fake reviews) is even more damaging as they can take total control of the sentiment on the target product due to its size. This work studies spam detection in the collaborative setting, i.e., to discover fake reviewer groups. The proposed method first uses a frequent item set mining method to find a set of candidate groups. It then uses several behavioural models derived from the collusion phenomenon among fake reviewers and relation models based on the relationships among groups, individual reviewers, and products they reviewed to detect fake reviewer groups.

Additionally, we also built a labelled dataset of fake reviewer groups. Although labelling individual fake reviews and reviewers is very hard, to our surprise labelling fake reviewer groups is much easier. We also note that the proposed technique departs from the traditional supervised learning approach for spam detection because of the inherent nature of our problem which makes the classic supervised learning approach less effective. Experimental results show that the proposed method outperforms multiple strong baselines including the state of supervised classification, regression, and learning to rank algorithms [12].

2.13 UNCOVERING COLLUSIVE SPAMMERS IN CHINESE REVIEW WEBSITES

C. Xu, J. Zhang, K. Chang, and C. Long, et.al ‘‘Uncovering collusive spammers in Chinese review websites, ‘In this work Community Question Answering (CQA) portals provide rich sources of information on a variety of topics.

However, the authenticity and quality of questions and answers (Q&as) has proven hard to control. In a troubling direction, the widespread growth of crowd sourcing websites has created a large-scale, potentially difficult-to-detect workforce to manipulate malicious contents in CQA. The crowd workers who join the same crowd sourcing task about promotion campaigns in CQA collusively manipulate deceptive Q&As for promoting a target (product or service). The collusive spamming group can fully control the sentiment of the target.

How to utilize the structure and the attributes for detecting manipulated Q&As? How to detect the collusive group and leverage the group information for the detection task? To shed light on these research questions, we propose a unified framework to tackle the challenge of detecting collusive spamming activities of CQA.

First, we interpret the questions and answers in CQA as two independent networks. Second, we detect collusive question groups and answer groups from these two networks respectively by measuring the similarity of the contents posted within a short duration. Second using attributes (individual level and group-level) and correlations (user-based and content based), we proposed a combined factor graph model to detect deceptive Q&As simultaneously by combining two independent factor graphs [13].

2.14 DISCOVERING SHILLING GROUPS IN A REAL E-COMMERCE PLATFORM

Y. Wang, Z. Wu, Z. Bu, J. Cao, and D. Yang et.al ‘‘Discovering shilling groups in a real e-commerce platform, ‘in this work As the use of recommender systems becomes generalized in society, the interest in varying the orientation of their recommendations is increasing.

There are shilling attacks' strategies that introduce malicious process in collaborative altering recommender systems in order to promote the own products or services or to discredit those of the competition. Academic research against shilling attacks has been focused in statistical approaches to detect the unusual patterns in user ratings.

Nowadays, there is a growing research area focused on the design of robust machine learning methods to neutralize the malicious process inserted into the system. This work proposes an innovative robust method, based on matrix factorization, to neutralize the shilling attacks. Our method obtains the reliability value associated with each prediction of a user to an item. By monitoring the unusual reliability variations in the items prediction, we can avoid promoting the shilling predictions to the erroneous recommendations [14].

2.15 PARTIALLY SUPERVISED LEARNING

B. Liu and W. S. Lee, ‘‘Partially supervised learning, ‘in this work We investigate the following problem: Given a set of documents of a particular topic or class , and a large set _ of mixed documents that contains documents from class and other types of documents, identify the documents from class.

The key feature of this problem is that there is no labelled non document, which makes traditional machine learning techniques inapplicable, as they all need labelled documents of both classes.

We call this problem partially supervised classification. In this work, we show that this problem can be posed as a constrained optimization problem and that under appropriate conditions; solutions to the constrained optimization problem will give good solutions to the partially supervised classification problem [15].

CHAPTER 3

MATERIALS AND METHODS

3.1 EXISTING SYSTEM

In existing work By conducting several public opinion surveys, based on their results it can be evaluated that people do read and get influenced by ratings and reviews of the products online. a partially supervised learning model (VWNB) to detect spammer groups. By labelling some spammer groups as positive instances, VWNB applies positive unlabeled learning (PU-learning) as a classifier to study positive instances from a spammer group detector (labelled spammer groups) and unlabeled instances (unlabeled groups). Specifically, we extract a reliable negative set the positive instances of the terms and the distinctive features.

By combining the positive instances, extracted Negative instances and unlabeled instances, we convert the PU-learning problem into a well-known semi supervised learning problem, and then use a Naive Bayesian model and an EM algorithm to train a classifier for spammer group detection survey performed by a leading site has shown that: More than 80% of the online customers look at the reviews available. 50% base their purchase on the ratings of the products. 30% of the customers compare the ratings of similar products before making their decision. Clearly consumers value the feedback given by other users as do the companies that sell such products. Blogs, websites, discussion boards etc. are a repository of customer suggestions which are valuable and important sources of textual data. Therefore, today's individuals and older ones extensively rely on reviews available on line. It means that people make their decisions of whether to purchase the products or not by analyzing and reflecting the existing opinions on those products.

The fact that is if the potential customer or users gets a genuine overall impression of a product by considering the present affect for that product, it is highly probable that he will actually purchase the product. Normally if the percentage of positive and effective opinions is considerable, it is likely that the overall impression will be highly positive. Likewise, if the overall impression is not proper, it is doubtful that they don't buy the product. Now the customers can write any opinion text, this can motivates the individuals, and organizations to give undeserving spam opinions to promote or not to credit some target products, services, organizations, individuals, and even ideas without disclosing their true intentions. These spammed opinion information is called opinion spam. VWNB Classification is a supervised machine learning technique. It is simple but one of the most effective techniques of classification but it takes too complex coding with poor functionality . It is an assumption based theorem. Even a little Violation of these assumptions will not affect the performance of the algorithm.

DRAWBACKS:

- ✓ Identifying reviews in the free text reviews, a straightforward solution is to employ an existing aspect identification approach.
- ✓ The spam classification instance grouping may low.
- ✓ Less accuracy prediction on opinion analysis**.**
- ✓ User review based word alignment is cumbersome.
- ✓ High in latency to analyze the datasets.
- ✓ Naive Bayes theorem is developed on the mathematical Bayes Theorem in probability increase ROC Value that reduces overall accuracy.

3.2 PROPOSED METHODOLOGY

In this proposed system the **Value Weighted Naïve Bayes with Frequent Pattern Ultra metric Tree** based opinion review analysis reviews possess the following characteristics: (a) they are frequently commented in user reviews; and (b) users' opinions on these reviews greatly influence their overall opinions on the reviews. A straightforward frequency-based solution is to regard the reviews that are frequently commented in user reviews as important. However, users' opinions on the frequent reviews may not influence their overall opinions on the reviews, and would not influence their purchasing decisions. We are measuring public concern using a two-step sentiment word alignment approach.

This work VWNB-FIUT Identifying fake reviews from a large dataset is challenging enough to become an important research problem. Business organizations, specialists and academics are battling to find the best system for opinion spam analysis. A single algorithm cannot solve all the problems' and challenges faced in today's generation with advancements in technologies, though a few are very efficient in analysis. It also improving the performance of the opinion spam analysis, and developing one that is consistently efficient across all categories of data.

The opinion reviews obtained from users can be classified into positive or negative reviews, which can be used by a consumer to select a product. This work aims to classify amazon reviews into groups of positive or negative polarity by using machine learning algorithms. In this study, we analyze online amazon reviews using proposed methods in order to detect fake reviews. The text classification methods are applied to a dataset of Amazon or google play store reviewa.

ADVANTAGES:

- ✓ The reviews containing explicit content and with swear words are not taken into consideration and are removed from the dataset.
- ✓ Sentiment score for each word is calculated when words are extracted into a form of dictionary or so called ‘Bag of Words (BOW)’ It first identifies the nouns and noun phrases in the documents. The occurrence frequencies of the nouns and noun phrases are counted, and only the frequent ones are kept as reviews.
- ✓ The language model was built on reviews, and used to predict the related scores of the candidate reviews. The candidates with low scores were then filtered out.
- ✓ The admin can easily identify related opinion reviews on that session.
- ✓ Easily determine reviews quality by using customer reviews.
- ✓ We can find Based on the number of reviews classified as Personal Negative; we compute a Measure of Concern (MOC) and a timeline of the MOC. We attempt to correlate peaks of the MOC timeline to peaks of the News (Non-Personal) timeline.
- ✓ Best accuracy results are achieved.
- ✓ Analysis of product after spam removal is done on the basis of their respective features.

MODULE DESCRIPTION

DOMAIN-SPECIFIC SENTIMENT KNOWLEDGE AND SPAM SIMILIARITY MODULE

A general observation in sentiment analysis field is that the words occurring more frequently in positive samples than negative samples in all domains convey positive sentiment orientations, and vice versa. More specifically, denote $s^{m} \epsilon R^{D*1}$as the sentiment word distribution of domain m, where $s_{w}^{m}$is the sentiment score of word w in this domain. Then $sm\ w$ is computed using word $w's$ association with positive sentiment label in domain m minus its association with negative sentiment label. In this module we use point wise mutual information (PMI) to measure the associations between words and sentiment labels, and the sentiment score of word $w$in domain w is formulated as

$$s_{w}^{m} = \mathrm{PMI}(w, posLabel_{m}) - \mathrm{PMI}(w, negLabel_{m})$$

$$= \log \frac{n\ (w,posLabel_{m})N_{m}}{n(w)n((w,posLabel_{m})} - \log \frac{n(w,negLabel_{m})N_{m}}{n(w)n((w,Label_{m})} \qquad (1)$$

$$= \log \frac{n(w, posLabel_{m})n(w, negLabel_{m})}{n(w, negLabel_{m})n(w, posLabel_{m})},$$

where $n(posLabel_{m})$ and $n(negLabel_{m})$ are the numbers of positive and negative samples in domain m respectively, and $n(w, posLabel_{m})$ and $n(w, negLabel_{m})$ are the frequencies of word w occurring in positive and negative samples of domain m. N_{m} is the number of all labeled samples in domain m. According to Eq. (1), if a sentiment word w has a higher probability to occur in positive samples than negative samples, then it will have a positive sentiment score, which indicates it tends to convey positive sentiment information in this domain.

For example, an unlabeled review in Kitchen domain is “the food processor is quick and easy for making baby food.” Since “quick” and “easy” are used to describe the same target in the same sentence, they probably convey the same sentiment. This is because people usually hold consistent opinions towards the same target in a short period, which is validated by social science theories such as sentiment consistency. If we can find more cases that these two words co-occur in the same sentence, then we can infer that they tend to convey similar sentiments in this domain. Thus, in this module we propose to extract domain-specific sentiment relations among words from the unlabeled samples based on their co-occurrence patterns.Denote S_{m} as the set of all sentences in the unlabeled samples of domain m. If a sentence contains adversative conjunctions such as “but” and “however”, we break it into clauses to make sure the sentiment expressed in each sentence of S_m is consistent. Then we count the frequencies of each word and each pair of words. Denote N_w^m as the frequency of word w, and $N_{w_1 w_2}^m$ as the frequency of word pair $\{w_1, w_2\}$ in domain m. Then the contextual similarity score between words $w1$ and $w2$ is formulated as follows:

$$C_{w_1 w_2}^m = \log \frac{N_{w_1 w_2}^m \cdot |S_m|}{N_{w1}^m . N_{w2}^m}, \qquad (2)$$

where$C_{w_1 w_2}^m$ represents the contextual similarity score between words $w_1$and w_2 in domain m, and $|S_m|$ is the size of S_m, i.e., the number of all sentences in the unlabeled samples of domain m. According to Eq. (2), if a pair of words has a higher probability to co-occur with each other in a specific domain, then they tend to share a higher contextual similarity score, which indicates that they probably convey similar sentiments in this domain. Note that the contextual similarity score defined in Eq. (2) can be negative. In this module, we only keep the positive similarities and filter out all the negative ones. Based on the contextual similarities among words extracted from massive unlabeled samples, we can build a domain-specific contextual similarity graph, where nodes represent words and edges represent the contextual similarities between words. In this module, we propose to use the domain-specific contextual similarity graph to refine the initial sentiment word distribution extracted from limited labeled samples to improve its accuracy and coverage. Denote $S^m \epsilon\ \mathbb{R}^{D*1}$as the initial sentiment word distribution of domain m, and $C^m \epsilon\ \mathbb{R}^{D*D}$as the domain-specific contextual similarity graph. Denote $P^m \epsilon\ \mathbb{R}^{D*1}$ as the final sentiment word distribution after refinement, which is computed according to a

$$\underset{pm}{argmin} \sum_{i=1}^{D} (p_i^m - s_i^m)^2 + \frac{\theta}{2} \sum_{i=1}^{D} \sum_{j \neq i} C_{i,j}^m (p_i^m - p_j^m)^2 \qquad (3)$$

$$= \|p^m - s^m\|_2^2 + \theta . (p^m)^T \mathrm{L^m p^m},$$

Where s_i^m the initial sentiment is score of the i_{th} word in domain m, and $C_{i,j}^m$is the contextual similarity score between the i_{th} and j_{th} words. $\mathrm{L}^m \epsilon\ \mathbb{R}^{D*D}$ Is the Laplacian matrix of C^m and $\mathrm{L}^m = \mathrm{D}^m - \mathrm{C}^m$ where D^m is a diagonal matrix and $D_{i,j}^m = \sum_{j=1}^{D} C_{i,j}^m$.The quadratic graph regularization in Eq. (3) is inspired by label propagation, and the inherent connection between label propagation and quadratic cost criterion was pointed by Bengio et al. in .

Since L^m in Eq. (3) is positive-semidefinite, the optimization problem in Eq. (3) is convex. We can derive that its analytical solution is $\mathrm{p}^m = (I + \theta \mathrm{L}^m)^{-1} \mathrm{s}^m$, where I ϵ $\mathbb{R}^{D*D}$ is an identity matrix. However, since matrix inversion is timeconsuming especially when dimension D is high, we propose to solve Eq. (3) using gradient descent method and use sparse matrix to store L^m.

SPAM SIMILIARITY TEXTUAL CONTENT BASED SPAM SIMILIARITY

The textual content based spam similarity is motivated by the observation that although different topics and opinion targets are discussed in different domains, similar domains may share many common terms. For example, in both Smart Phone and Digital Camera domains, terms like "screen", "battery", and "image" are frequently used. In contrast, the probability of two far different domains such as Smart Phone and Book sharing many common terms is low. Thus, we propose to measure the similarity between domains based on their textual content. Inspired by the work in, here we select Jensen Shannon divergence to measure the similarity of two domains based on their textual term distributions. Denote d^m ϵ $\mathbb{R}^{D*1}$ and d^n ϵ $\mathbb{R}^{D*1}$ as the term distribution vectors of domains m and n respectively, where D represents the dictionary size. $d\,_t^m \epsilon$ [0,1] stands for the probability of term *t* occurring in domain m. Then the textual content based SPAM SIMILIARITY between domains *m* and *n* is formulated as

$$ContentSim_{(m,n)} = 1 - D_{JS}(\ \mathrm{d}^m \| \mathrm{d}^n)$$

$$= 1 - \tfrac{1}{2}\Big(D_{KL}\big(\mathrm{d}^m \| \bar{\mathrm{d}}\big) + D_{KL}\big(\mathrm{d}^n \| \bar{\mathrm{d}}\big)\Big), \qquad (4)$$

where$\bar{\mathrm{d}} = \frac{1}{2}(\mathrm{d}^m + \mathrm{d}^n)$ is the average distribution, $D_{JS}(.)$ represents Jensen-Shannon divergence, and $D_{KL}(.)$ is the Kullback-Leibler divergence which is defined as:

$$D_{KL}(\mathrm{p}\|\mathrm{q}) = \textstyle\sum_{i=1}^{D} p(t) \log_2 \frac{p(t)}{q(t)}. \qquad (5)$$

Since the base of logarithm used in Eq. (5) is 2, $D_{JS}(\ \mathrm{d}^m \| \mathrm{d}^n)\ \epsilon\ [0,1]$. Thus, the range of the textual content based SPAM SIMILIARITY defined in Eq. (4) is also [0,1].

SENTIMENT EXPRESSION BASED SPAM SIMILIARITY

The textual content based Spam Similarity introduced in previous section can measure whether two domains have similar word usage patterns. However, high similarity in textual content does not necessarily mean that sentiment words are used in similar ways in these domains. For example, both CPU and Battery belong to electronic hardware. In CPU domain, the word “fast” is usually positive. For instance, “Intel Core i7 is very fast.” However, in Battery domain, the word “fast” is frequently used as a negative word (e.g., “This battery runs out too fast”). Thus, measuring SPAM SIMILIARITY based on sentiment expressions may be more suitable for multi-domain spam review classification task.

Denote p^m and p^n as the sentiment word distributions of domains m and n respectively, which are extracted from both labeled and unlabeled samples according to previous section. Then the sentiment expression based SPAM SIMILIARITY between domains m and n is defined as the cosine similarity of their sentiment word distributions

$$Sentisim_{(m,n)} = \frac{p^m \cdot p^n}{\|p^m\|_2 \cdot \|p^n\|_2}. \qquad (6)$$

Note that $Sentisim_{(m,n)}$ defined in Eq. (6) can be negative in theory, although the probability is very small. In this module, we constrain that domain similarities should be nonnegative. Thus, if the SentiSim score between a pair of domains is negative, then we set it to zero.

SIMILARITY ANALYSIS AND SPAM REVIEW CLASSIFICATION

First, we introduce several notations that will be used in following discussions. Assume there are M domains to be analyzed. Denote$\{ X^m \epsilon \mathbb{R}^{N_m * D}, Y^m \epsilon \mathbb{R}^{N_m * D}\}$ as the labeled samples in domain$m$, where $N_m$is the number of labeled samples and D is the size of feature set. $X_i^m \epsilon \mathbb{R}^{D*1}$ is the transpose of the ith row of X^m, standing for the feature vector of the ith labeled sample in domain m, and y_i^m is its sentiment label. In this module we focus on binary spam review classification, i.e., classifying a piece of text into positive or negative, and $y_i^m \epsilon \{+1, -1\}$. Denote the global sentiment classifier shared by all domains as $w \ \epsilon \ \mathbb{R}^{D*1}$, and the domain-specific sentiment classifier of domain m as $w_m \epsilon \ \mathbb{R}^{D*1}$. We use matrix $W \ \epsilon \ \mathbb{R}^{D*M}$, to represent all domain-specific sentiment classifiers, where the *mth* column of W stands for the domain-specific spam review classification model of domain m, i.e.,$W_{.,m} =$

w_m. Denote the loss of classifying X_i^m into label y_i^m under parameters w and W as $f(X_i^m, y_i^m, w + W_{.,m})$, where f is classification loss function, which can be squared loss, log loss, hinge loss and so on.

In this module, we propose to extract prior general sentiment knowledge from general-purpose sentiment lexicons to enhance the learning of the global spam review classification model$w$. Denote $P \epsilon \mathbb{R}^{D*1}$ as the general sentiment knowledge extracted from an existing sentiment lexicon. If word i is included in the sentiment lexicon and labeled as positive (or negative), then $p_i = 1(or - 1)$ Otherwise,$p_i = 0$. In addition, denote $P^m \epsilon \mathbb{R}^{D*1}$ as the domain-specific sentiment knowledge extracted from both labeled and unlabeled data of domain m, and p_i^m is the prior sentiment score of word i. Besides, denote $S \epsilon \mathbb{R}^{M*M}$ as the SPAM SIMILIARITY matrix. $S_{m,n} \epsilon [0,1]$ Is the similarity between domains m and n. The diagonal elements of S are set to zeros.

Model

Given multiple domains to be analyzed, a small number of labeled samples in these domains, the domain similarities between them, the general sentiment knowledge extracted from general-purpose sentiment lexicons, and the domain-specific sentiment knowledge of each domain extracted from both labeled and unlabeled samples, the goal of our approach is to train accurate domain-specific sentiment classifiers for multiple domains in a collaborative way. The model of our proposed approach is formulated as an optimization problem as follows:

$$\begin{aligned} arg \min_{w,W} \mathcal{L}(w, W) = \sum_{m=1}^{M} \sum\nolimits_{i=1}^{N_m} f(x_i^m, y_i^m, w + W_{.,m}) - \propto_1 w^T P \\ - \propto_1 \textstyle\sum_{m=1}^{M} (w + W_{.,m})^T P^m + \beta \textstyle\sum_{n \neq m}^{M} S_{m,n} \left\| W_{.,m} - W_{.,m} \right\|_2^2 \\ + \lambda_1 (\|w\|_2^2 + \|W\|_F^2) + \lambda_2 \left(\|w\|_1 - \|W\|_{1,1} \right), \end{aligned} \tag{7}$$

where $\propto_1$,$\propto_2$, and β are non-negative regularization coefficients for general sentiment knowledge, domain-specific sentiment knowledge, and SPAM SIMILIARITY knowledge respectively. λ_1 and $\lambda_2$are positive coefficients of the i_2- and ‘ i_1-norm regularization terms. $\|w\|_F$ and $\|w\|_{1,1}$ are the Frobenius norm and $l_{1,1}$-norm of matrix W respectively. $\|W\|_F = \sqrt{\sum_{i=1}^{D} \sum_{j=1}^{D} W_{i,j}^2}$ and $\|W\|_{1,1} = \sum_{i=1}^{D} \sum_{j=1}^{M} \left| W_{i,j} \right|$ Inspired by the idea of Lasso, we introduce the l_1-norm of the global spam review classification model and the $l_{1,1}$-norm of the domain-specific spam review classification models to control the sparsity of these models. Since not all words convey sentiment information, introducing the l_1-norm of model parameters can be regarded as conducting feature selection for sentiment words. We also combine the l_1-norm regularization terms with the l_2-norm regularization terms motivated by elastic net regularization, which can improve model stability in high-dimensional problems.

According to Eq. (7), the final sentiment classifier of each domain is a linear combination of the global spam review classification model and the domain-specific model of this domain. The global spam review classification model is shared by all domains, and is trained on the labeled samples from all domains. The domain-specific spam review classification model is trained using the labeled data within one domain. By splitting the sentiment classifier of each domain into a global component and a domain-specific component, we can better model the general sentiment knowledge shared by various domains and at the same time capture the domain-specific sentiment knowledge more accurately. Following many previous works in sentiment analysis area here we select linear classification model in our approach, which has good spam review classification performance and excellent interpretability. Various classification loss functions can be selected for f in our model. For example, f can be squared loss (i.e., $f\left(x_i^m, y_i^m, w + W_{.,m}\right) = \left(\left(w + W_{.,m}\right)^T . x_i^m - y_i^m)^2\right)$,hinge loss (i.e., $f\left(x_i^m, y_i^m, w + W_{.,m}\right) = \left[1 - y_i^m\left(w + W_{.,m}\right)^T \cdot x_i^m\right]_+$),and logloss (i.e., $f\left(x_i^m, y_i^m, w + W_{.,m}\right) = \log(1 + \exp\left(-y_i^m\left(w + W_{.,m}\right)^T \cdot x_i^m\right)))$.Minimizing the term $\sum_{m=1}^{M} \sum_{i=1}^{N_m} f\left(x_i^m, y_i^m, w + W_{.,m}\right)$ in Eq. (7) means that we hope the sentiment classifiers learned by our approach can classify the labeled samples in each domain as accurately as possible. In this way, we incorporate the sentiment information in labeled samples into the sentiment classifier learning.

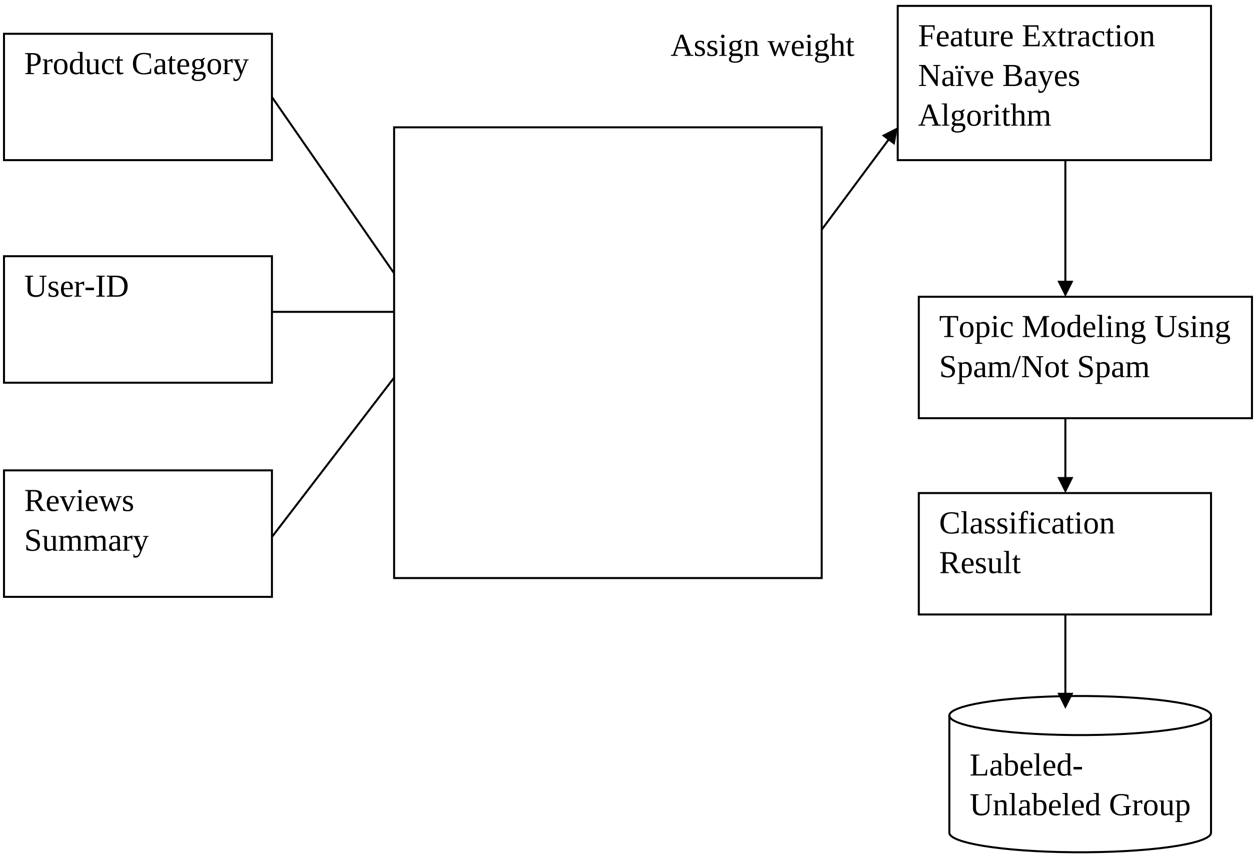
In Eq. (7), by the term $-\propto_1 w^T P$ we constrain that the global spam review classification model learned by our approach is consistent with the prior general sentiment knowledge extracted from sentiment lexicons. For example, since "excellent" is a positive sentiment word in many existing sentiment lexicons such as SentiWordNet, we expect that the sentiment score of "excellent" in our global spam review classification model is also positive. If its sentiment score is negative in our model, then a penalty will be added to the objective function. In this way, the prior general sentiment knowledge extracted from general-purpose sentiment lexicons can be used to guide the learning of the global spam review classification model in our approach. Similarly, through the term $-\propto_2 \sum_{m=1}^{M}\left(w + W_{.,m}\right)^T P^m$ we incorporate the domain-specific sentiment knowledge into sentiment classifier training. If a word w has a positive (or negative) prior sentiment score in the domain-specific sentiment knowledge of domain m, then we hope its sentiment score in the final sentiment classifier of this domain is also positive (or negative).

From Eq. (7), the SPAM SIMILIARITY knowledge is incorporated into our model as graph regularization of the domain-specific spam review classification models. Minimizing the term $\sum_{m=1}^{M} \sum_{n\neq m}\left\|W_{.,m} - W_{.,n}\right\|_2^2$ means that if two domains share a high SPAM SIMILIARITY with each other, then the sentiment classifiers of these two domains should be more similar. In this way, we encourage similar domains to share more sentiment information with each other than dissimilar domains.

RATING PREDICTION MODULE

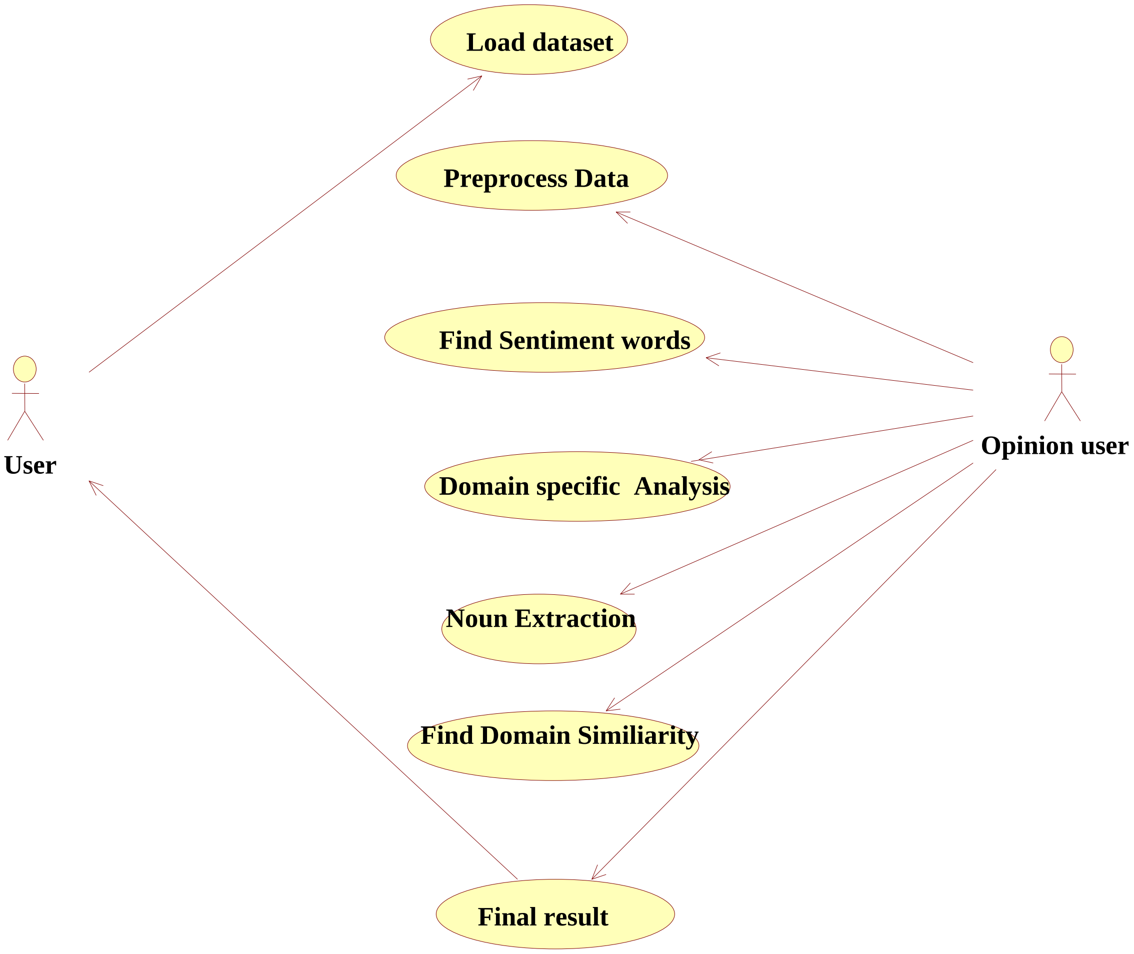
Since the objective function in our model (Eq. (7)) is convex, learning the optimal sentiment classifiers in our approach is equivalent to solving a convex optimization problem. However, even if the loss function f is smooth, the optimization problem in Eq. (7) is still nonsmooth due to the l_1--norm regularization of the global spam review classification model and the $l_{1,1}$-norm regularization of the domain-specific spam review classification models. In addition, the learning processes of sentiment classifiers in different domains are coupled together in our approach in order to exploit the sentiment relatedness among these domains. Thus, it is challenging to solve the optimization problem in our approach efficiently. In this module, we introduce an accelerated algorithm based on VWNB to solve the model of our approach. In addition, we propose a parallel algorithm based on FIUT to train sentiment classifiers for multiple domains in a parallel way, which can further improve the efficiency of our approach when domains to be analyzed are massive. Next we will introduce them in detail.

OVERALL ARCHITECTURE DIAGRAM

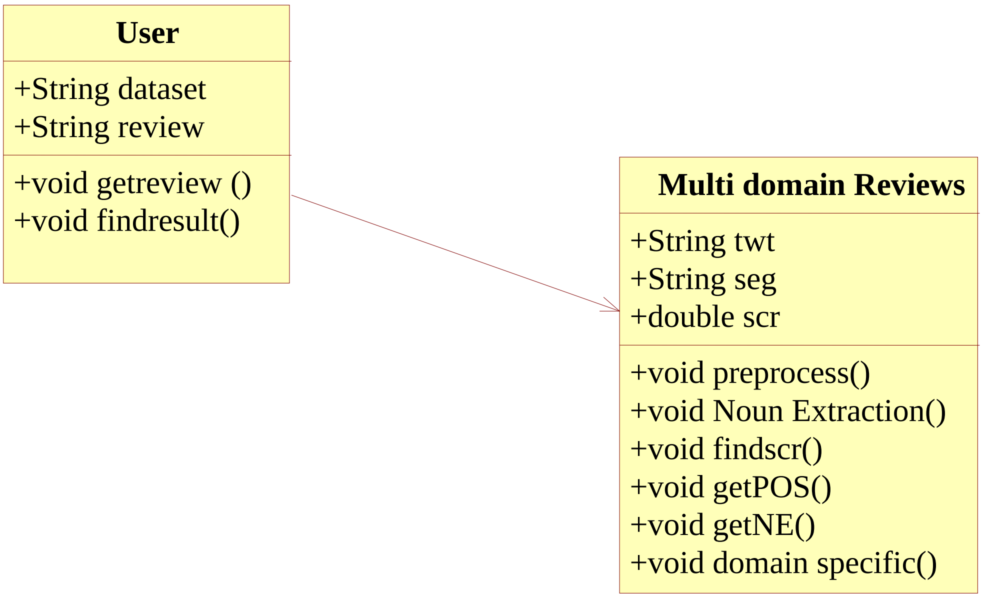


UML Diagrams

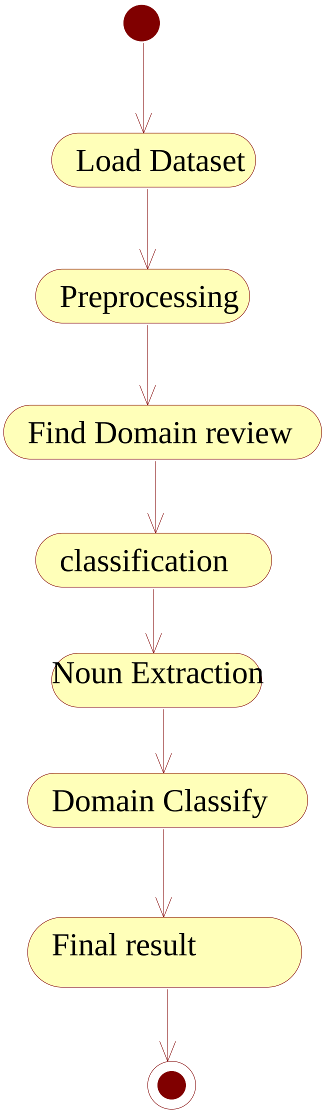
Use Case Diagram



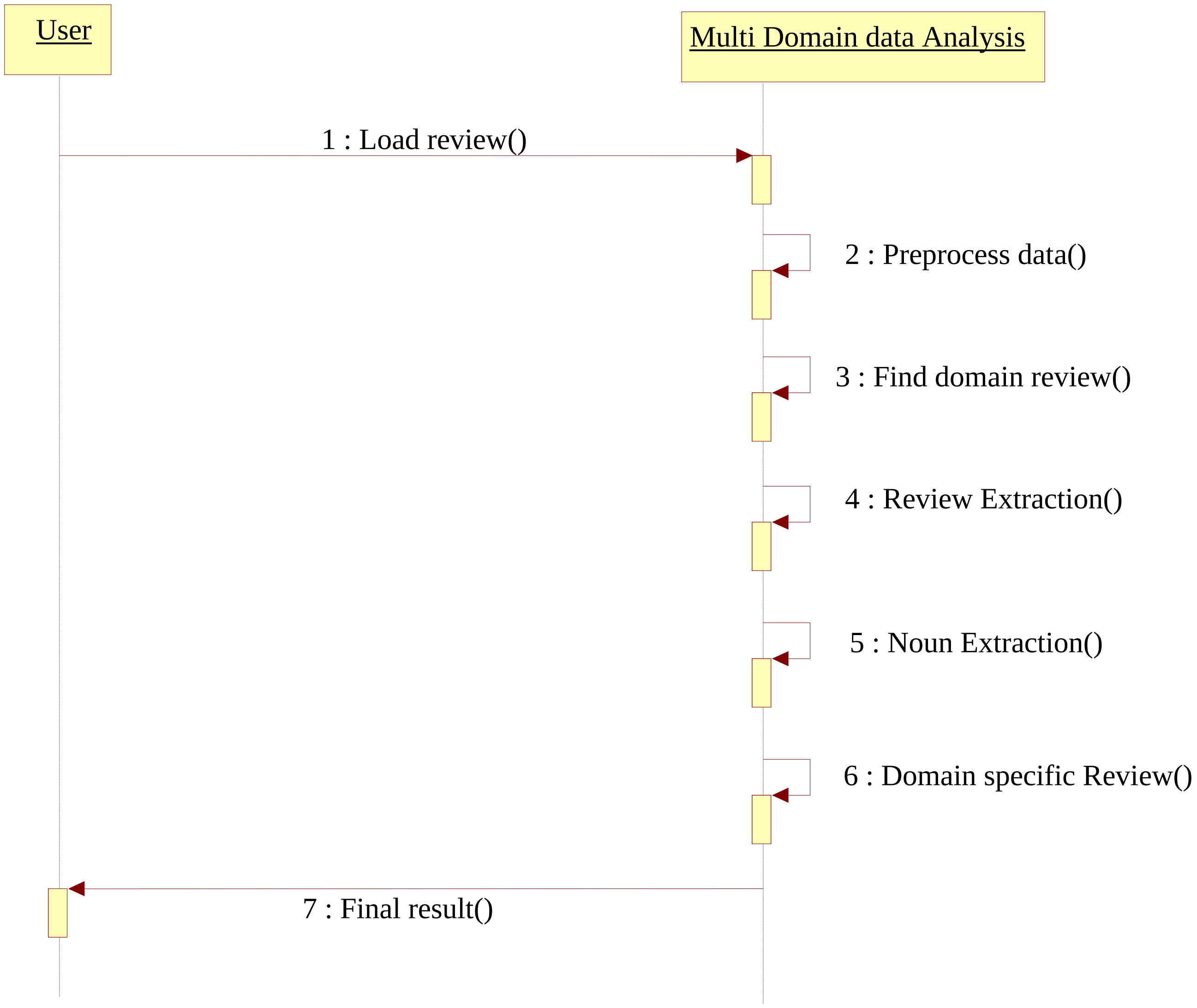
CLASS DIAGRAM



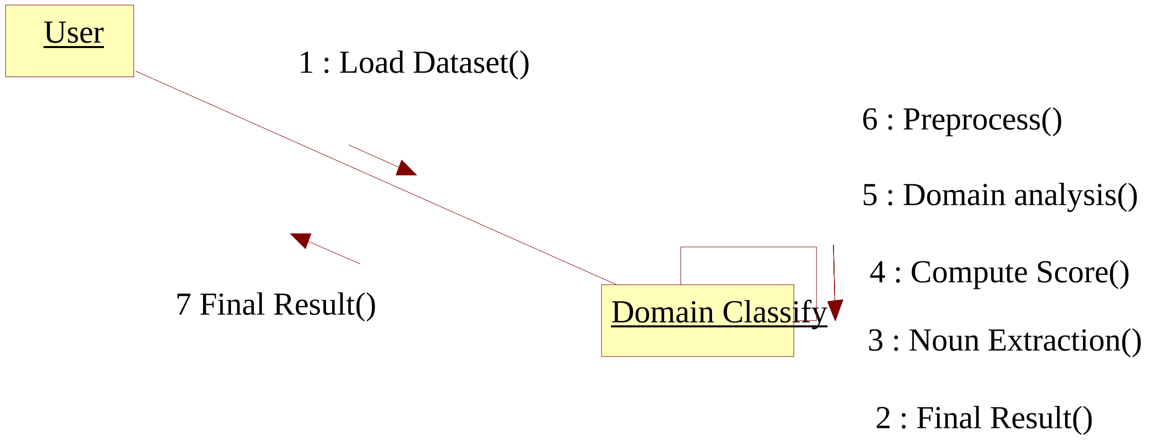
ACTIVITY DIAGRAM



SEQUENCE DIAGRAM



COLLABORATION DIAGRAM



CHAPTER 4

EXPERIMENTAL SETUP AND PROCEDURE

The experiment with the real-world data was performed to check the credibility and reliability of Twitter system with positive results. Both TSP and SS filtering were proposed by using partial data for real time and lightweight spammer detection. Both algorithms contain some false positive, but their true positive are not better to collusion rank. A hybrid approach that uses attributes of both filtering are suggested. The experiment was performed on thousand authorized users and thousand spammer accounts with social status and TSP features. The result of the proposed approach shows that the schemes are scalable because they check user cantered two hops social network instead of examining the whole network.This study significantly improves the performance of false and true positives than the previous scheme.

CHAPTER 5

RESULT AND DISCUSSION

we dissected that malignant exercises on online media are being acted in a few different ways. In addition, the scientists have endeavoured to recognize spammers or spontaneous bloggers by proposing different arrangements. Accordingly, to consolidate every appropriate exertion, we proposed a scientific categorization as per the extraction and classification strategies. The classification depends on different classifications, for example, counterfeit content, URL based, moving points, and by recognizing counterfeit clients. The first significant classification in the scientific categorization is of strategies proposed for recognizing spam, which is infused in the Twitter stage through phony substance. Spammers for the most part consolidate spam information with a point or catchphrases that are malignant or contain the sort of words that are probably going to be spam

CONCLUSION

This work proposes a partially supervised learning based Model VWNB to detect spammer groups from product reviews. First, the frequent item mining using the VWNB model (FIM) to discover spammer group candidates from the review data. Then, manually labeling some spammer groups as positive instances, the VWNB employs construct to PU-Learning the positive and unlabeled instances of a classier to identify the real candidates from the group of real spammer groups. In particular, the VWNB dense a feature strength function The group features of the measure of discriminatory power, and then High discriminative with iteratively removes instances Get a reliable set of unlabeled instances from only non-spammer groups of negative set consisting. By combining the positive, negative and unlabeled instances, we The well-known semi supervised into the PU-Learning problem learning problem, and employer Naive Bayesian Model and EM algorithm to construct a classified as spammer group detector. Experiments on Amazon.cn show that the proposed VWNB model outperforms both supervised and Spammer group detection on unsupervised learning methods. Improvement in the area of our future work of the VWNB model.

Beyond the Naive Bayesian model used in VWNB, we will investigate and incorporate more classification models such as neural network, Semi-Supervised SVM (S3VM) and even ensemble methods. On the positive instances acquisition and RN extraction, we plan to involve Improving the accuracy and efficiency of active learning data labeling.

REFERENCES

[1] F. Zhu and X. Zhang, ‘‘Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics,’’ J. Marketing, vol. 74, no. 2, pp. 133–148, 2010.

[2] K. C. Santosh and A. Mukherjee, ‘‘on the temporal dynamics of opinion spamming: Case studies on yelp,’’ in Proc. 25th Int. Conf. World Wide Web, 2016, pp. 369–379.

[3] N. Jindal and B. Liu, ‘‘Opinion spam and analysis,’’ in Proc. Int. Conf. Web Search Data Mining, 2008, pp. 219–230.

[4] F. Li, M. Huang, Y. Yang, and X. Zhu, ‘‘Learning to identify review spam,’’ in Proc. Int. Joint Conf. Artif. Intell. (IJCAI), 2011, vol. 22. no. 3, pp. 219–230.

[5] S. Xie, G. Wang, S. Lin, and P. S. Yu, ‘‘Review spam detection via temporal pattern discovery,’’ in Proc. 18th ACM SIGKDD Int. Conf. Know. Discovery Data Mining, 2012, pp. 823–831.

[6] A. Mukherjee et al., ‘‘Spotting opinion spammers using behavioral footprints,’’ in Proc. 19th ACM SIGKDD Int. Conf. Know. Discovery Data Mining, 2013, pp. 632–640.

[7] L. Akoglu, R. Chandy, and C. Faloutsos, ‘‘Opinion fraud detection in online reviews by network effects,’’ in Proc. ICWSM, vol. 13. 2013, pp. 2–11.

[8] S. Shehnepoor, M. Salehi, R. Farahbakhsh, and N. Crespi, ‘‘Net Spam: A network-based spam detection framework for reviews in online social media,’’ IEEE Trans. Inf. Forensics Security, vol. 12, no. 7, pp. 1585–1595, Jul. 2017.

[9] Z. Wu, Y. Wang, Y. Wang, J. Wu, J. Cao, and L. Zhang, ‘‘Spammers detection from product reviews: A hybrid model,’’ in Proc. IEEE Int. Conf. Data Mining (ICDM), Nov. 2015, pp. 1039–1044.

[10] A. Heydari, M. Tavakoli, and N. Salim, ''Detection of fake opinions using time series,'' Expert Syst. Appl., vol. 58, pp. 83–92, Oct. 2016.

[11] H. Li et al., ''Bimodal distribution and co-bursting in review spam detection,'' in Proc. 26th Int. Conf. World Wide Web, 2017, pp. 1063–1072.

[12] A. Mukherjee, B. Liu, and N. Glance, ''Spotting fake reviewer groups in consumer reviews,'' in Proc. 21st Int. Conf. World Wide Web, 2012, pp. 191–200.

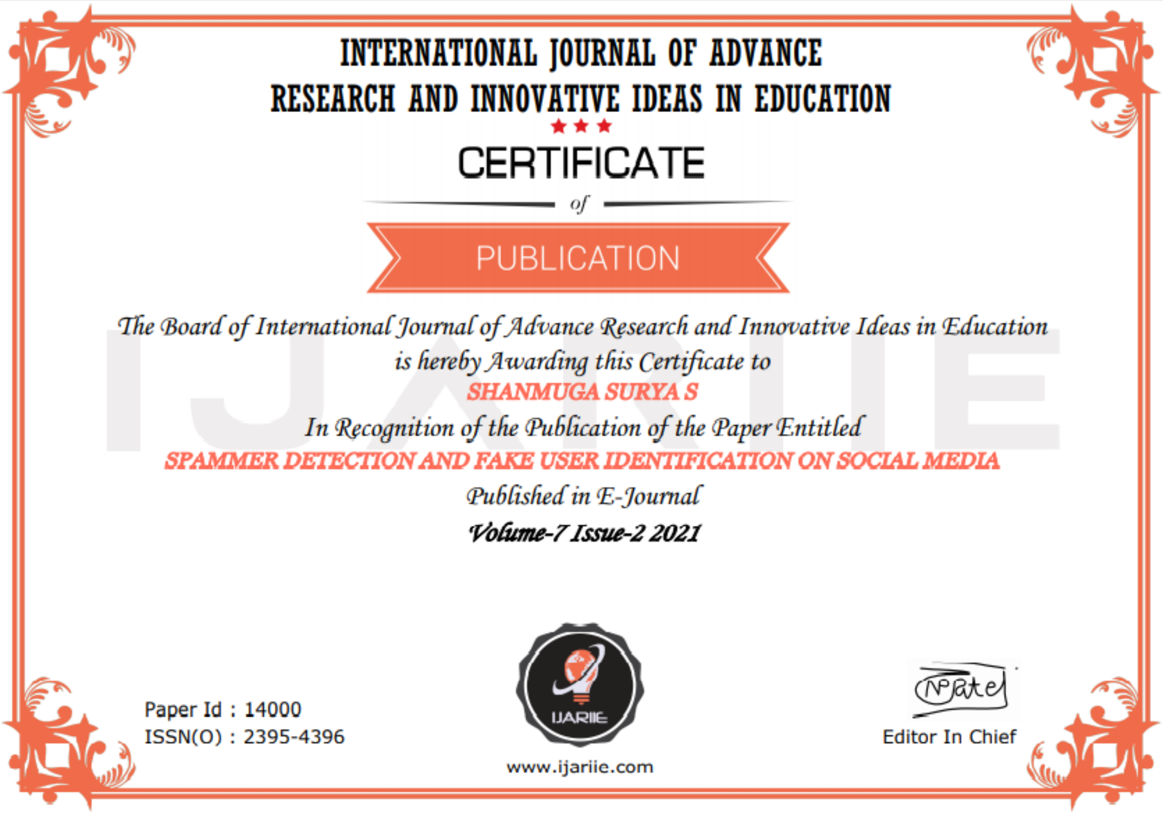
[13] C. Xu, J. Zhang, K. Chang, and C. Long, ''Uncovering collusive spammers in Chinese review websites,'' in Proc. 22nd ACM Int. Conf. Conf. Inf. Know. Manage. 2013, pp. 979–988.

[14] Y. Wang, Z. Wu, Z. Bu, J. Cao, and D. Yang, ''Discovering shilling groups in a real e-commerce platform,'' Online Inf. Rev., vol. 40, no. 1, pp. 62–78, 2016.

[15] B. Liu and W. S. Lee, ''partially supervised learning,'' in Web Data Mining. Berlin, Germany: Springer, 2011, pp. 171–208.

LIST OF PUBLICATIONS

[1] Srihariharan T, Shanmuga Surya S, and Senthilraja P. "Spammer Detection And Fake User Identification on Social Media" International Journal Of Advance Research And Innovative Ideas In Education Volume 7 Issue 2 2021 Page 1150-1153



INTERNATIONAL JOURNAL OF ADVANCE
RESEARCH AND INNOVATIVE IDEAS IN EDUCATION

★★★

CERTIFICATE

of

PUBLICATION

The Board of International Journal of Advance Research and Innovative Ideas in Education is hereby Awarding this Certificate to

SHANMUGA SURYA S

In Recognition of the Publication of the Paper Entitled

SPAMMER DETECTION AND FAKE USER IDENTIFICATION ON SOCIAL MEDIA

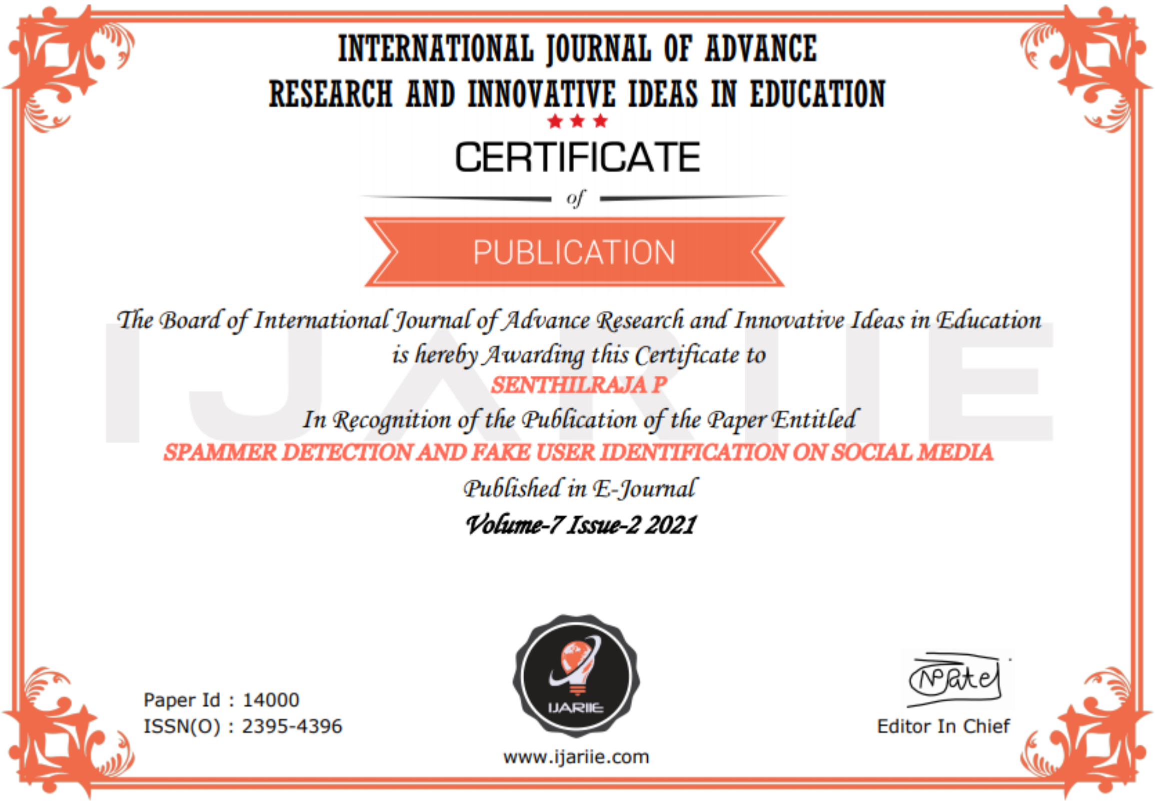
Published in E-Journal

Volume-7 Issue-2 2021

Paper Id : 14000
ISSN(O) : 2395-4396

IJARIIE
www.ijariie.com

NPatel
Editor In Chief



INTERNATIONAL JOURNAL OF ADVANCE RESEARCH AND INNOVATIVE IDEAS IN EDUCATION

★★★

CERTIFICATE

of

PUBLICATION

The Board of International Journal of Advance Research and Innovative Ideas in Education is hereby Awarding this Certificate to

SENTHILRAJA P

In Recognition of the Publication of the Paper Entitled

SPAMMER DETECTION AND FAKE USER IDENTIFICATION ON SOCIAL MEDIA

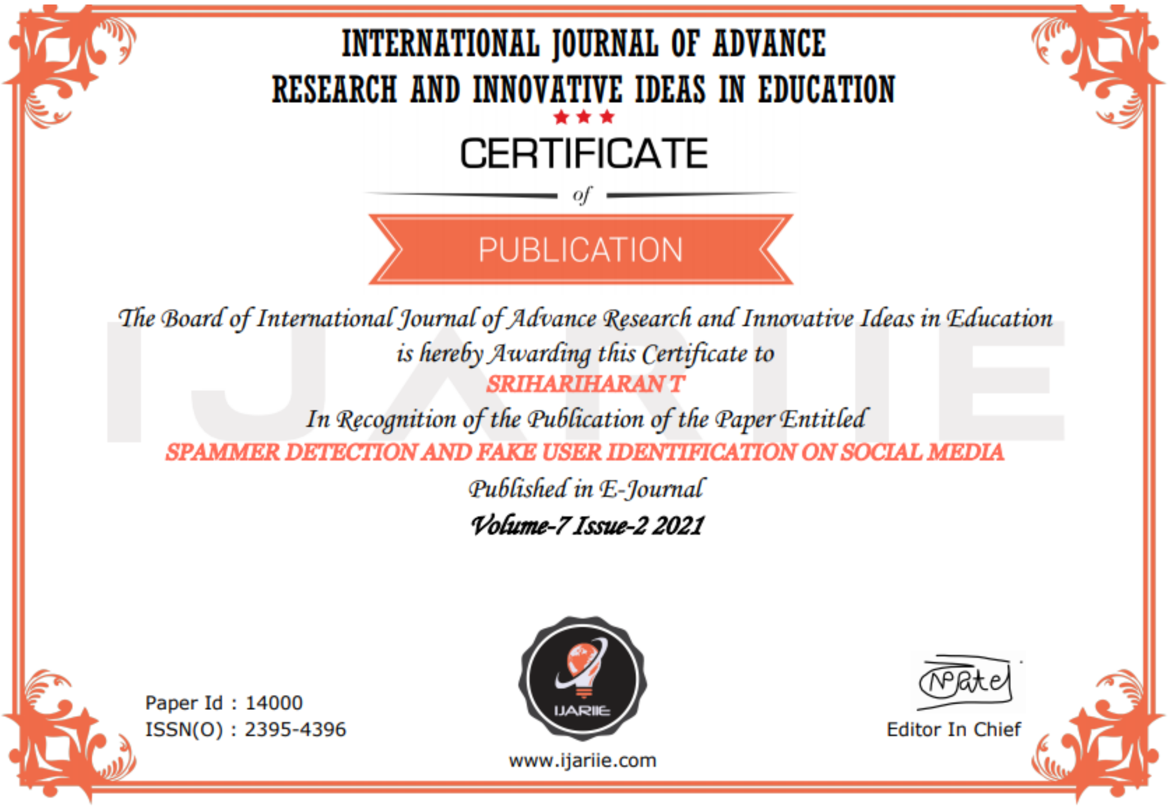
Published in E-Journal

Volume-7 Issue-2 2021

Paper Id : 14000
ISSN(O) : 2395-4396

www.ijariie.com

NPatel
Editor In Chief



INTERNATIONAL JOURNAL OF ADVANCE RESEARCH AND INNOVATIVE IDEAS IN EDUCATION

★★★

CERTIFICATE

of

PUBLICATION

The Board of International Journal of Advance Research and Innovative Ideas in Education is hereby Awarding this Certificate to

SRIHARIHARAN T

In Recognition of the Publication of the Paper Entitled

SPAMMER DETECTION AND FAKE USER IDENTIFICATION ON SOCIAL MEDIA

Published in E-Journal

Volume-7 Issue-2 2021

Paper Id : 14000
ISSN(O) : 2395-4396

www.ijariie.com

NPatel
Editor In Chief

SPAMMER DETECTION AND FAKE USER IDENTIFICATION ON SOCIAL MEDIA

SRIHARIHARAN T[1], SHANMUGASURYA S[2], SENTHILRAJA P[3]

1, Student, Department of Computer Science and Engineering, K.S. Rangasamy College of Technology, Tiruchengode, Tamil Nadu

2, Student, Department of Computer Science and Engineering, K.S. Rangasamy College of Technology, Tiruchengode, Tamil Nadu

3, Assistant Professor, Department of Computer Science and Engineering, K.S. Rangasamy College of Technology, Tiruchengode, Tamil Nadu

ABSTRACT

As e-commerce is growing and becoming popular day-by-day, the number of reviews received from customer about any product grows rapidly. People nowadays heavily rely on reviews before buying anything. Product reviews play an important role in deciding the sale of a particular product on the ecommerce websites or applications like Flipkart, Amazon, Snapdeal, etc. In this paper, we propose a framework to detect fake product reviews or spam reviews by using Opinion Mining. The Opinion mining is also known as Sentiment Analysis. In sentiment analysis, we try to figure out the opinion of a customer through a piece of text. The proposed method called VWNB-FIUT (Value Weighted Naïve Bayes with Frequent Pattern Ultra Metric Tree) automatically classifies users' reviews into "suspicious", "clear" and "hazy" categories by phase-wise processing. The hazy category recursively eliminates elements into suspicious or clear. This results into richer detection and be useful to business organization as well as to customers. Business organization can monitor their product selling by analyzing and understanding what the customers are saying about products. This can help customers to purchase valuable product and spend their money on quality products. Finally end users see that each individual review with polarity scores and credibility score annotated on it. We first take the review and check if the review is related to the specific product with the help of VWNB. We use Spam dictionary to identify the spam words in the reviews by using FIUT. In Text Mining we apply several algorithms and on the basis of these algorithms we get the specific results.