ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 3, Issue 11, November 2016

Automated Content Based Short Text Classification forFiltering Undesired Posts on Facebook

[1] Geethika K [2] Jyothi Tiwari P [3] Monika S [4] Ganapriya R [5] Asst . Prof. P Chandana M
[1][2][3][4]Department of Computer Science and Engineering,
Sri Sairam College of Engineering, Bengaluru.

***Abstract:* -- Online Social Networking (OSN) sites are always helpfulforbeingsocializedand togetexposedtoasocial environment. But, privacy and prevention of undesired posts on user wallis the only problem of biggest concern. User should have the ability to control the message posted on their own private wall to avoid undesirable contents to be displayed. The existing OSN sites have very little support regarding this problem. For example, Facebook filters messages on the basis of identity of senderi.e. only friend, friend of friend or group of friends can post any message; no content based preferences are supported. Taking this fact into consideration, the proposed work contributes to address such problem through a machine learning baseds of the classifie for labeling messages in support of contents of message. This work experimentally evaluates an automated scheme to filter out unwanted messages posted on Facebook walls by as signing a set of categories with each short text message based on its contents.**

***IndexTerms*—Online Social Networkingsites,OSN,Content based classification, Short Text Classifier, Machine Learning, wall posts,facebookwall,user profile**

I. INTRODUCTION

Social networking site sare the mostinter active medium for sharing information like; photos,blogs, thoughts, reviews, comments etc. This social media is gaining much importance and popularity day by day over most of the other mediums for business perspective such as; advertising various sells products, promoting new mobile applications and many more. Most of the social networking sites have common features in them. The basic feature of social networking sites is the ability to create and share a personal profile. Social networks also let you post photos, statuses,feelings and personal belongs on your profile page. One of the most common features of online social networks is to find and make friends over the network. These friends also appear as links, so visitors can easily browse your online friend network. According to face book statistics,1 million links are shared, 2 million friends are requested and 8-10 million messages have been send for every 30 minutes on Facebook.

The content present in social network is constituted by short text, and the not able example is the messages written by Social Network users on particular private or public areas, known as general walls. In support of displaying contents of user's own wall most of the social sites has the feature of preventing the messages from unwilling people on the basis of friendship status. In the presents cenario, let us suppose if any one of the user's friend posted some objection able text, before user get the notification and before removing that message from time line manually it might have been seen by many other users, which should not happen. So,there should be a mechanism which will automatically restrict such posts. Up till nowno feature is present which filters the message in accordance with the contents of the message, no matter about the identity of the person who is posting it. Taking this problem into consideration, an automated system, to filter undesired comments from owner's wall is proposed. For filtering the short length text messages onuser wall content based short text classification methodology is proposed here.

Generally, content based filtering algorithms are mostly used in recommender systems for calculating the utility values for particular item and recommending other items to the user which have higher degree of similarity to the user's profile. Where in, here the content based filtering technique is applied for the social networking site Facebook for filtering unwanted messages or posts on userwall.

The proposed system includes, the machine learning based short text classifier, a social network manger and the content based message filtering rules. According to the survey on related work it has been found that many authors proposed systems which are based on contents based filtering, but no one had worked on actual existing online social networking sites. The proposed work is the only one which is actually implemented on existing online social networking site Facebook.

There mainder of the paper is organized as follows: section II gives the problem definition, section III informs about the related work,section IV, V, VI gives the brief over view of the Methodology, Basic Architecture and experimental design. Finally,sectionVII concludesthepaper.

II.PROBLEMDEFINITION

Currently, OSN does not provide message content based preferences to control messages on userwall. Therefore, it is not possible to prevent undesired messages, such as political or vulgarones, without concerning about the other user who posts them. As the wall messages consists of short text with limited word occurrences and include informal / colloquial abbreviated language; providing as ervice of message filtering is not only amatter of using web content mining techniques, rather it requires to design ad-hoc short text classification strategies, Hence, the automated content-based short text classification technique I required to filter out unwanted messages.

III.RELATEDWORK

Content based filtering is generally used for recommender system sorf or web page filtering. The basic idea of content based filtering came through authors Gediminas Adomavicius and Alexander Tuzhilin[2].Their paper presents an overview of the field of recommender systems and describes the current generationof recommendationmethodsthatareusually classified into the following threemain categories: content-based, collaborative, and hybrid recommendation approaches.

AuthorHuiLi, Fei Caiand Zhi fang Liao [3]have combined probabilistic model and classical content-based filtering recommendational gorithms to proposea new algorithm for recommendation system, using Hidden Markov Model. The basic approachd escribed in this paper is calculating the similarity of user profile and each profile of all the items and recommending item to satisfy user need or tastes.

Michael Chauand Hsinchun Chen [4]expanded the idea of content based filtering for filtering the webpages. They used the ML paradigm along with Web content analysis and Web structure analysis. Marco Vanetti, Elisabetta Binaghi, Moreno Carullo,Elena Ferrari and Barbara Carminati [1]put for ward the idea of using content based filtering for OSN. They provide the facility to have straight rules over their own wall to avoid the unwanted messages to be posted.The basic aim of having a control over the posts is achieved through a Filtered wall(FW). The system described in this paper blocks the undesired messages sent by the user. But the drawback of this system is that, the user will not be blocked; only the message posted by the user will be blocked. To over come this problem, Black list rule can be implemented as future enhancement.

The messages on facebook consist of short texts. Handling such short text messages for filtering purpose is one of the majorissues becauses hort texts do not have sufficient word occurrences. To deal with such short text messages authors Josh Weissbock, Ahmed A.Esmin, Diana Inkpen[5] proposed the methodology to enhance the text of the messages that contain link with externalin for mation such as the title of the webpages and/or most frequent terms from these webpages. This paper also says that the results of the classification improve substantially by adding this externalin formation.

Authors Sriram, Bharath[6] again proposed the approach which effectively classifies the text to apre defined set of generic classessuch as News, Events, Opinions, Deals, and Private Messages on the basis of author information and features with in the twits. Accordingtosurvey on relatedworkonthe sameconcept, many authors proposed their work. Reddy, M.Vamsi Krishna [7]proposed their work on the same content filtering principle. They created a web application which performs the functionality, but it is restricted for only one computer as it is a standal one application where admin has rights to generate filteringpolicies.

Rose,J.Anishya and A.Pravin [8],Ezhilvani,V.,K. Malathiet.al[9], Dhruv VashisthaandSivagami.G.[10] ,Bala Kumar,Bercelin Rose Mary and Devi Mareeswari[15] are many other authors who proposed their work which defines the filtering and black list rules for filtering the posts. Authors Thilagavathi,N., and R.Taarika[13] proposed their work on same line for filtering messages posted on OSNwalls. They used the inference algorithm to in fer the new rules from the existing rules in support of content based filtering. From all this related survey we inferred that, efforts have been taken for filtering undesired posts on OSN created by them.But,tillnow no work has been done on the actual Online Social Networking site.They all worked on OSN prototype. Hence, we proposed the scheme which will be directly implemented on actual facebook site.

IV.METHODOLOGY

It has been found that noone worked on the actual existing social networking sites like twitter or Facebook. Hence,by considering the further enhancement, application of filtering rules on real time Facebook site has been invented. The proposed work is the implementation of the filtering rules on the contents posted on Facebook user walls to avoid unwanted posts t obedisplayed. The techniques used are explained here shortly,

A.Content BasedFiltering

Information filtering systemsare designed to classifya stream of dynamically generated information dispatched asynchronouslybyaninformationproducer and present to the user those information that are likely to satisfy his/her requirements. Incontent-based filtering, each user is assumed to operate independently. As are sult, acontent - based filtering system selects information items based on the correlation between the content of the items and the user preferences as opposed to a collaborative filtering system that chooses items based on the correlation between people with similar preferences.

B.Short text Classifier

The message sposted on OSN walls are usually short text messages. Handling such short text for filtering purpose is one of the major issues because short texts do not have sufficient word occurrences. For such situations, a short text classification technique is applied here to supports hort text messages. The basic aim of using the short text classifier is to recognize and remove the neutral sentences and to categorize them. This classifier is hierarchical [fig1].It consists of two levels. The first level is the hard classifier level, in which message are classified with neutra landnon-neutrallabels.

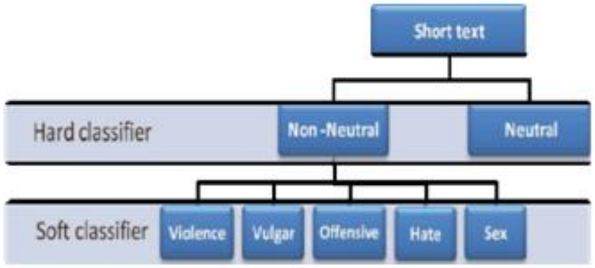

Fig.1. TextClassifier

In the secondlevel which is the soft classifier level, all non- neutral messages develop gradual membership. These grades will be used in succeeding phases for filtering process by considering proper there shold values. Short text classifier includes text representation and Machine Learning based Classification.

C.Text Representation[1]

Extracting the particular features from a given set of information is crucial task; this affects the entire performance of classification. For text representation the module used is Vector space model (VSM). According to VSM the document is represented as a vector of binary or real weights.

Dj= w1j…w|T|j

Where ,گ is a set of terms/features that occurat least once in at least one document of the collection گ rand wkjא [0;1] represent show much term tk contributes to the semantics of document dj.

In the case of non-binary weighting, the weight wkj of term tk (here,term represents word) indocument dj is computed according to the standard term frequency -- inverse document frequency (tf--idf) weighting function[1], defined as, tf–idf(tk, dj)=#(tk, dj).log (1) Where #(tk ,dj)denotes thenumber of times tk occur in document dj and # Ҁr(tk) de notes the document frequency of term tk. Document property is adopted by collecting the correct words, bad words, capital words, punctuation character, exclamation marksetc.In moredetail,

- Correct Words: This expresses the number of terms tk אگ̇اK, where tk is the term of document dj and K denotes the known words.

- Bad Words: This can be computed same as correct words feature, here K de note the dirty words in particular language.

- Capital Words: This denotes the amount of words writtenincapitalletters.Itiscalculatedas fractionof words having more than half of the characters in capitalcase. Fore.g. Valueof"To DO orNoT TOdo" is 0.5; excludingthe initial word of message.

- Punctuation Characters: This can be calculated by the fraction of total number of punctuation character in the message by total number of characters in the message. For e.g. the value of this feature for the message "hi!! What's up?"is5/14.

- Exclamation Marks: This can be calculated as the fraction of the total numbers of exclamation marks in the message by total number of punctuation characters in the message. For e.g. Consider "Hello!!! How r u?" is 3/4.

- Question Marks: This can be calculated as the fraction of total numbers of question mark in the message by total number of punctuation character in the message. For e.g. consider "hi! How u doing?"is 1/2.

D.Machine Learning Based Classification Technique[1] Consider M1 and M2 be the two levels of classifier. Let vectory1 be the belonging ness to the neutral class. Suppose, TrS D and Te SD are training and test sets respectively.

The learning and generalization phases work as follows:

- For meach message mi,if feature vector xii sextracted, the trainingandtestsetsarethentransformedas,TrS=
- {(xi;yi),…(x|TrSD|,y|TrSD|)} andTeS={(xi;yi),…(x|TeSD|,
- y|TeSD|)}.Binaryset iscreatedformessageM1 as TrS1= {(xj,yj)
- =jy ,(jy,jx)|SrTא yj1}.ForM2amulticlasstrainingsetiscreatedas,TrS2 =
- {(xj, yj)אTrS|(xj, yj|),yjk=yjk+1,k=2,…,|ŏ|}.M1istrainedwithTrS1to recognizewhethermessage isneutralornon-neutraland evaluatedusingTeS1.M2istrainedwiththenon-neutralTrS2forcomputing gradualmembership to the non-neutral classes. The testingisdonethroughTeS2.

V.BASICARCHITECTURE

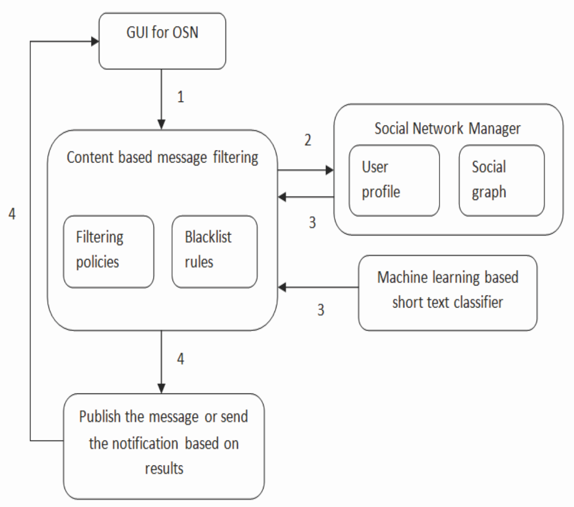
GUI for OSN
1
Content based message filtering
Filtering policies
Blacklist rules
2
3
Social Network Manager
User profile
Social graph
Machine learning based short text classifier
3
4
4
Publish the message or send the notification based on results

Fig.2. Basic Architecture

- When user wants to post a message on a private wall, he/she tries to enter into a wall and the user tries to post messagesontheprivatewallbutitisinterceptedbythe filter wall.(1)
- Secondlymachinelearningtextclassifieris usedtoextract themetadatafromthegivencontentofmessages.(2)
- Thenthefilterwallmakeuseofthismetadatawhichis providedbytheshorttextclassifierandalong withthe data extractedfrom theuser profilebyenforcingthefiltering rule.(3)
- Thefilterwallspublish or blockthemessagedepending on the result ofprevious step.(4)

A.Filtering Rules

For defining the filtering rules the issue that shouldbe considered may be themessagewith differentmeaning and significancebasedonthecreatorofmessage.Hence, here the type, depth, and trust value are recognized by creator specification.

Definition 1: (Creator specification) A Creator specification creatorSpec implicitly denotes a set of OSN users. Itcanhave one of thefollowingforms,possibly combined:

A set of attribute constraints of the forman OP av, where an is a user profile attribute name, av and OP are, respectively, a profile attribute value and a comparis on operator,compatible with an's domain. A set of relationship constraints of the form(m,rt, min Depth, max Trust), denoting all the OSN users participating with user mina relationship of typert, having a depth greater than or equal to min Depth,and a trust value less than or equal to max Trust.

Example 1: The creator specification CS1=(Age<16, Sex=male) denotes all the males whose age is less than 16 years, where as the creator specification CS2 = (Henry, colleague,2,0.4) denotes all the users who are colleagues of Henry and who set rust level is less than or equal to 0.4. Finally, the creator specification CS3 = (Henry,colleague,2,0.4,(Sex = male)) selects only the maleusers from those identified by CS2.

The final component of a FR is the action that the system has to perform whether block or notify, with the obvious semantics of blocking the message, or notifying the wall owner. An FR is there fore formally defined as follows:

Definition 2: (Filteringrule) A filtering rule FR is at uple (author, creator Spec, content Spec, action),where, author is the user who specifies the rule. Creator Spec is a creator specification, specified according to

Definition 1. Content Spec is a Boole an expression defined on content constraints of the form (C,ml) where Cisa class of the first or second level andml is the minimum membership level there shold required for class C to make the constraint satisfied. action ,kcolb)א notify) denotes the action to be performed by the system on the messages matching content Spec and created by users identified by creator Spec. In the proposed work above explained techniques are experimentally evaluatedon Facebook.

VI.EXPERIMENTALDESIGN

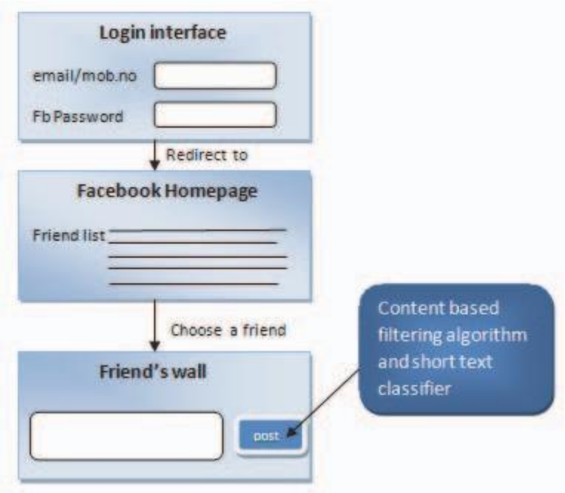

Fig.3. ExperimentalFlow

There isoneloginterface where userhastoenterhis/her logindetailsforsigninginto Facebook account.Thisredirects theusertothe Facebook homepage.Usercanselectany other usertosendmessage.Whenusertriestosendamessageit first get interrupted by the content based filteringalgorithm. Basedonthe results generated byalgorithmtheobjectionable post isrestricted.Ifthemessagecontentis neutralthesystem will post the message.

VII.CONCLUSION

Traditionally,contentbasedfiltering wasusedfor recommender systems.But,here we used thecontentbased filteringtechniquealongwithMachinelearningbasedshort textclassifiertofilterunwanted messages postedon Facebookuserwalls.Thisleadstothepreventionofundesired poststo bedisplayedtootherusers. However,infuturewecanapply thesetechniqueson othersocial networkingapplications like Twitter,Whatsapp etc.

REFERENCES

[1]MarcoVanetti,ElisabettaBinaghi,ElenaFerrari,Barbara CarminatiandMoreno Carullo,“asystemtofilterunwanted messagesfromOSNuserwalls”,IEEETrans.Knowledgeand Data Eng.,vol. 25,no.2,Feb2013.

[2]A.AdomaviciusandG.Tuzhilin,“TowardtheNextGeneration of RecommenderSystems: ASurveyof theState-of-the-Artand PossibleExtensions”,IEEETrans. KnowledgeandDataEng., vol.17,no. 6, pp.734-749,June 2005.

[3] Li,Hui,FeiCai,andZhifangLiao,"Content-BasedFiltering RecommendationAlgorithmUsingHMM", ComputationalandInformationSciences(ICCIS),2012Fourth International Conferenceon. IEEE, 2012

[4] Chau, Michael, and Hsinchun Chen, "A machine learning approachtowebpagefiltering usingcontentandstructureanalysis",DecisionSupportSystems44.2(2008):482-494.

[5] Weissbock,Josh,AhmedEsmin,andDianaInkpen, "Using External Information for Classifying Tweets", Intelligent Systems(BRACIS),2013Brazilian Conferenceon.IEEE,2013.

[6]Sriram,Bharath,etal,"Shorttextclassificationintwitterto improve informationfiltering",Proceedingsofthe33rd internationalACMSIGIRconferenceon Researchand developmentininformationretrieval.ACM,2010.

[7]Reddy,M.VamsiKrishna,etal,"ContentBasedFilteringin SocialNetworking SitesUsingWebApllication”,International journalof Soft Computingand EngineeringISSN:2231-2307, volume-4,May-2014.

[8]Rose,J.Anishya, andA.Pravin,"MachineLearningText CategorizationinOSN toFilterUnwantedMessages”, Internationaljournalof ComputerScienceandInformation Technologies,Vol.5(1),2014.

[9]Ezhilvani,V.,K.Malathi,andR.Nedunchelian,"RuleBased MessageFilteringAnd BlacklistManagementForOnlineSocialNetwork”,IJRET, Volume:03Issue: 06,Jun-2014

[10]Dhruva Vashstha and Sivagami G., “Filtering Undesired MessagesfromOnlineSocialNetworks:A ContentBased Filtering Approach”, (IJCSIT) International Journal of ComputerScienceandInformationTechnologies,Vol.5(2), 2014

[11]Apt, Chidanand, Fred Damerau, and Sholom M. Weiss. "Automatedlearningofdecisionrulesfor textcategorization."ACMTransactionsonInformationSystems(TOIS)12.3(1994):233-251.

[12] Dumais, Susan, et al. "Inductive learning algorithms and representations for text categorization." Proceedings of the