Unwanted Message Filtering in Online Social networks (OSN) using Machine Learning Technique

R. Rajeshwari[1] & Mrs. Rajapriya D[2]

[1]*Research Scholar, Department of Computer Science, RVS Technical Campus (Affiliated to Anna University), Coimbatore-641402, Tamil Nadu, India.*

[2]*Professor, Department of Computer Science, RVS Technical campus (Affiliated to Anna University), Coimbatore-641402, Tamil Nadu, India.*

Abstract: *Online Social networks (OSNs) are extremely popular collaboration and communication tools that have attracted millions of Internet users. Unfortunately, recent evidence show that they can also be affective mechanism for spreading attacks. Popular OSNs are increasingly becoming the target of phishing attacks launched from large botnets. One basic issue in today's Online Social Network (OSNs) is to give users the power to regulate the message denote on their own non-public house to avoid that unwanted content to displayed. Two recent studies have confirmed that the existence of large-scale spam campaigns in Twitter and Face book, respectively. Up to now, OSNs offer very little support to the current demand. To fill the gap, proposed system permitting OSN users to own an on the spot management on the messages denote on their walls. This is often achieved through an adaptable rule-based system that permits users to customize the filtering criteria to be applied to their walls, and a Machine Learning-based soft classifier mechanically labelling messages in support of content-based filtering.*

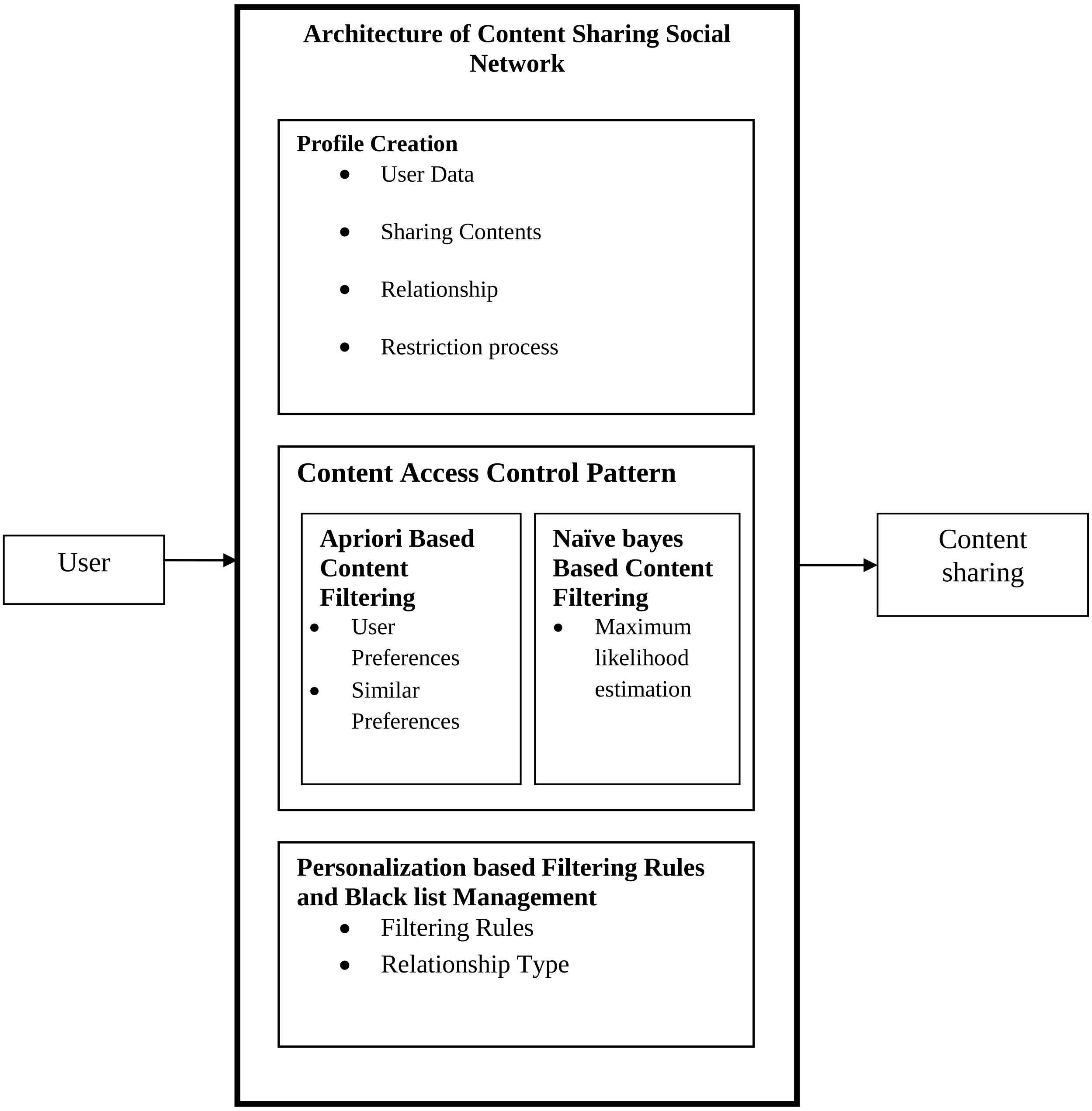
1. INTRODUCTION

One basic issue in today's Online Social Networks (OSNs) is to give users the power to regulate the messages denote on their own non-public house to avoid that unwanted content is displayed. Up to now, OSNs offer very little support to the current demand. To fill the gap, proposed system permitting OSN users to own an on the spot management on the messages denote on their walls. This is often achieved through an adaptable rule-based system that permits users to customize the filtering criteria to be applied to their walls, and a Machine Learning-based soft classifier mechanically labeling messages in support of content-based filtering. In the proposed system, we apply novel machine learning algorithm for text classification based on a Naive Bayes classifiers. Results obtained show that Naive Bayes classifier algorithm has good accuracy then the existing Apriori algorithm and is thus comparable to any other methods for unwanted message classification

Web-based services are used to extract the significant information from large quantity of data respectively. For example Facebook is the most popular social networking site in which millions of people have opened their user account. Facebook provides all type of services like adding friends, recommending friends, sharing of images, audio and video etc. But Facebook also provides facility to user to post the message on users wall. So, there is possibility that posted message could be vulgar or offensive one. Which may cause serious problems like harassing or blackmailing can also happen, it means instead of all those advantages there are some disadvantages with Social networking sites. Thus, we can say that OSNs provide poor security to user. To avoid such types of problems we can use Information filtering. Information filtering is nothing but it checks whether the content of message is pleasant or not. If the content of message is impolite then Information filtering technique will not allow the person to post on wall.

2. PROBLEM STATEMENT

Today OSNs provide very little support to prevent unwanted messages on user walls. For example, Facebook allows users to state who is allowed to insert messages in their walls (i.e., friends, friends of friends, or defined groups of friends). However, no content-based preferences are supported and therefore it is not possible to prevent undesired messages, such as political or vulgar ones, no matter of the user who posts them. However, no content-based preferences are supported and therefore it is not possible to prevent undesired messages, such as political or vulgar ones, no matter of the user who posts them. Providing this service is not only a matter of using previously defined web content mining techniques for a different application, rather it requires to design ad hoc classification strategies



METHODS

LIST OF MODULES

- Construction of Online Social Network Architecture
- Construction of Online Social Network Dashboard
- Construction of Filtered OSN User wall
- Construction of Short Text Classifier using Apriori & Naive Bayes Algorithm
- Performance comparison

Construction of Online Social Network Architecture

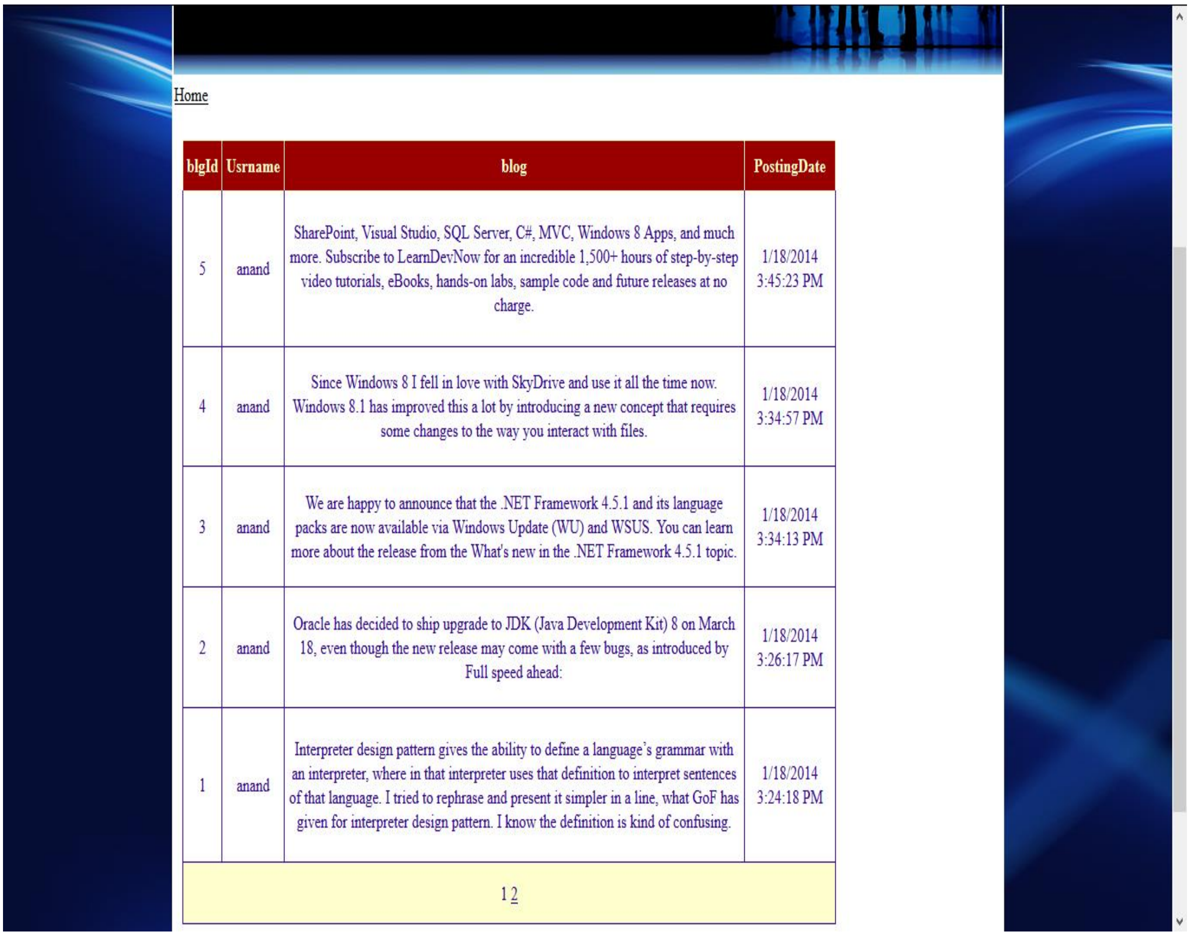
We leverage Face book as the running example in our discussion since it is currently the most popular and representative social network provider. In the meantime, we reiterate that our discussion could be easily extended to other existing social network platforms. To provide meaningful and attractive services, these social applications consume user profile attributes, such as name, birthday, activities, interests, and so on. To make matters more complicated, social applications on current OSN platforms can also consume the profile attributes of a user's friends.

Construction of Online Social Network Dashboard:

In this module administrator is having the rights to manage the sub modules like Stop word dictionary, Correct words dictionary, Bad words dictionary, Punctuations characters, Classes and Terms etc.,.

Construction of Filtered OSN User wall:

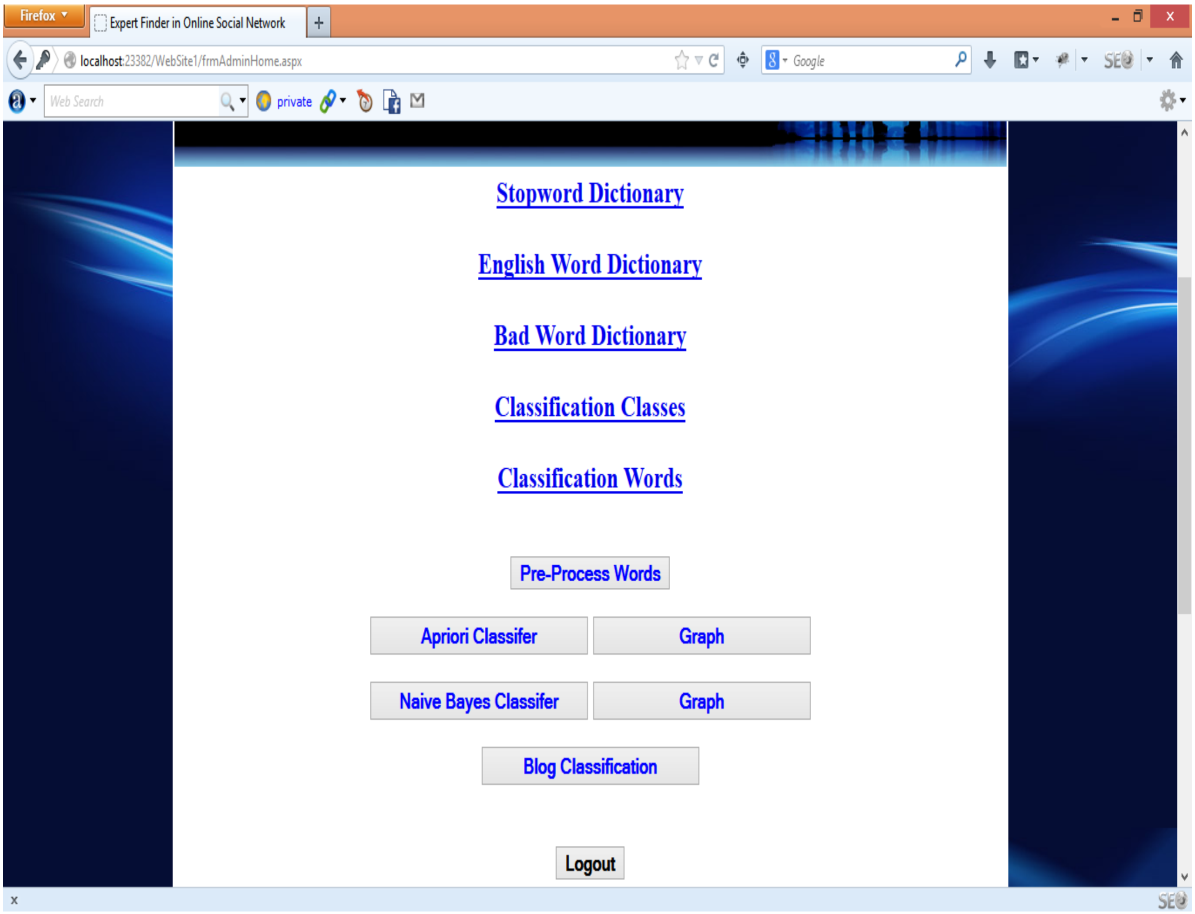
In this module, the Content-Based Messages Filtering architecture is configured. After entering the private wall of one of his/her associates, the user attempts to post a message, which is captured by Filtered Wall. A Machine Learning based text classifier extracts the data from the content of the message. Filtered Wall uses data provided by the classifier, mutually with data extorted from the social graph and users' profiles, to implement the filtering and Block List rules.



Construction of Short Text Classifier using Apriori & Naive Bayes Algorithm:

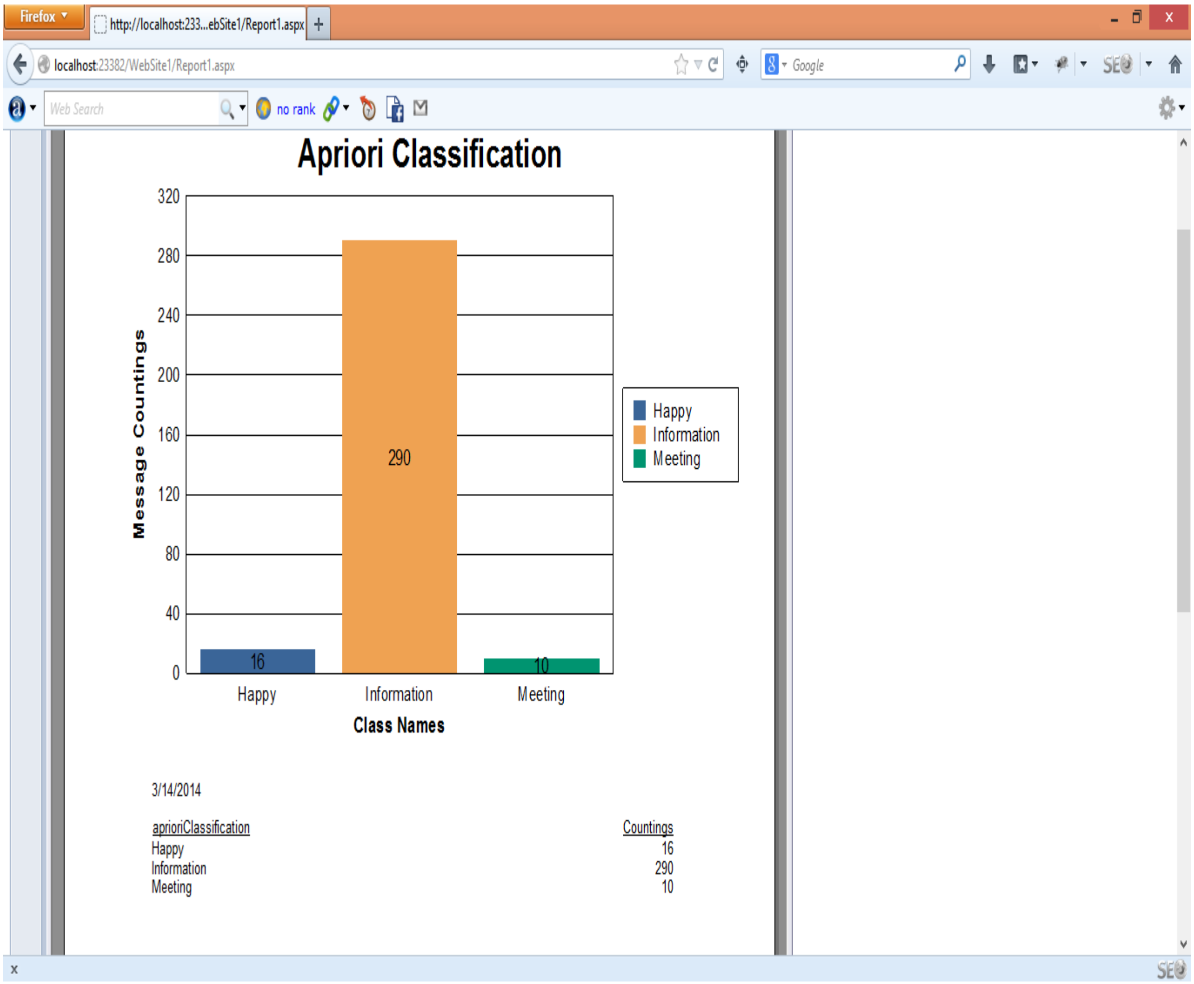
In this context, critical aspects are the definition of a set of characterizing and discriminate features allowing the representation of underlying concepts and the collection of a complete and consistent set of supervised examples. Our study is aimed at designing and evaluating various representation techniques in combination with a neural learning strategy to semantically categorize short texts. From a ML point of view, we approach the task by defining a hierarchical two-level strategy assuming that it is better to identify and eliminate "neutral" sentences, then classify "non-neutral" sentences by the class of interest instead of doing everything in one step. The first-level task is conceived as a hard classification in which short texts are labeled with crisp Neutral and Non-neutral labels. The second-level soft classifier acts on the crisp set of non-neutral short texts and, for each of them, it "simply" produces estimated appropriateness or "gradual membership" for each of the conceived classes, without taking any "hard" decision on any of them. Such a list of grades is then used by the subsequent phases of the filtering process.

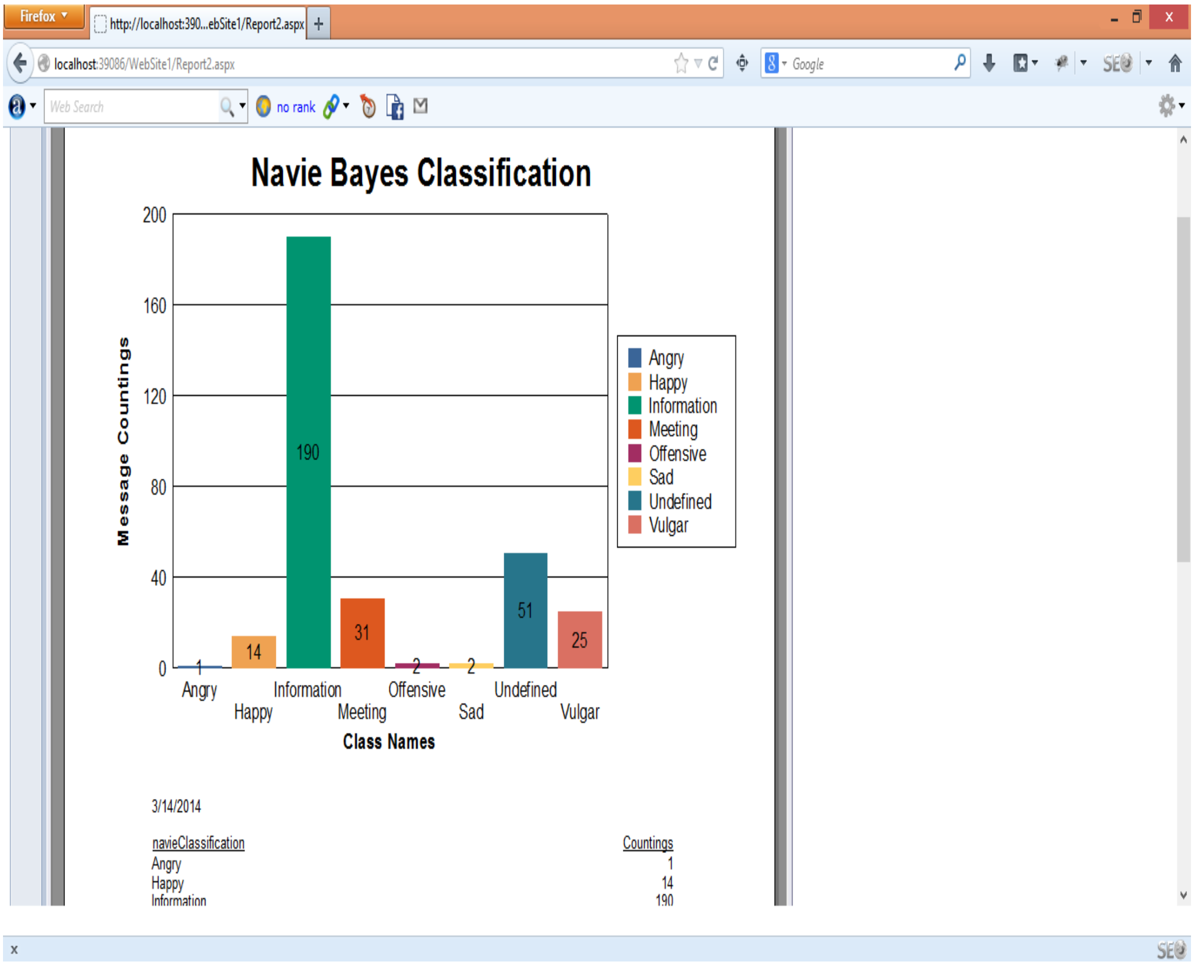
In this module, Apriori algorithm is designed to operate on databases containing transactions. Other algorithms are designed for finding association rules in data having no transactions or having no timestamps. The algorithm attempts to find subsets which are common to at least a minimum number C of the item sets. Apriori uses a "bottom up" approach, where frequent words of the messages are extended one item at a time (a step known as candidate generation), and groups of candidates are tested against the data. The algorithm terminates when no further successful extensions are found. Apriori, while historically significant, suffers from a number of inefficiencies or trade-offs, which have spawned other algorithms. Candidate generation generates large numbers of subsets. Length of rule can be limited by user defined threshold (MST). With smaller item sets the interpretation of rules is more intuitive. Unfortunately this can increase the amount of rules too much. Due to this drawback of Apriori algorithm and to improve the accuracy of classification we using Naive Bayes algorithm. In Naive Bayes algorithm, it is a classification algorithm based on Bayes rule, that assumes all the attributes X1,…,Xn are conditionally and mutually independent given Y. The value of this assumption dramatically simplifies and reduces the complexity and representation of P(X|Y) and the problem of estimating it from the training data. Calculating the final probability of unwanted message we can finally make the decision of being Neutral or Non-neutral depending on their majority value.



Performance comparison:

A huge number of short texts are generated every day, which calls for a method that can efficiently accommodate new data to incrementally adjust classification models. We apply smoothing models namely Apriori and Naive Bayes for unwanted message classification and study their performance. The experimental results on a message dataset show that the smoothing methods are able to significantly improve the unwanted message classification performance of Naive Bayes better than Apriori.





DATABASE DESIGN

Database design is the process of producing a detailed data model of a database. This logical data model contains all the needed logical and physical design choices and physical storage parameters needed to generate a design in a Data Definition Language (DDL), which can be used to create a database. A fully attributed data model contains detailed attributes for each entity.

TABLE BADWORDS

Column name	Data Type	Allow nulls
Badword	Varchar(200)	Not null

TABLE CLASS NAMES

Column name	Data Type	Allow nulls
Class names	Varchar(200)	Not null

TABLE CLASS TERMS

Column name	Data Type	Allow nulls
Class names	Varchar(200)	Not null
Engwords	Varchar(200)	Not null

TABLE FRIEND REQUEST

Column name	Data Type	Allow nulls
Usrname	Varchar(50)	Not null
Frndname	Varchar(50)	Not null
Frndreqstatu	Varchar(10)	Not null

TABLE PROFILE

Column name	Data Type	Allow nulls
Usrname	Varchar(50)	Not null
Usrbook	Varchar(500)	Not null
Usrmovie	Varchar(500)	Not null
Usrpolitical	Varchar(500)	Not null
Usrreligion	Varchar(500)	Not null
Usractivities	Varchar(500)	Not null
Usrsports	Varchar(500)	Not null
Usrsongs	Varchar(500)	Not null
usrtvshows	Varchar(500)	Not null
usrmobilemodel	Varchar(500)	Not null
usrmobilenetwork	Varchar(500)	Not null

TABLE SHORT TEXT

Column name	Data Type	Allow nulls
recId	Int	Not null
Shorttext	Varchar(50)	Not null
Fullword	Varchar(500)	Not null

TABLE STOP WORDS

Column name	Data Type	Allow nulls
Stword	Varchar(200)	Not null

TABLE USERBLOG

Column name	Data Type	Allow nulls
BlgId	Int	Not null
UsRNamE	Varchar50)	Not null
Blog	vaRcHar(500)	Notnull
Posting date	Atetime	nOtnull

TABLE USER REGISTRATION

Column name	Data Type	Allow nulls
Usrname	Varchar(50)	Not null
Paswrd	Varchar(50)	Not null
Regname	Varchar(50)	Not null
Usremail	Varchar(50)	Not null
Cntyname	Varchar(50)	Not null
Ctyname	Varchar(50)	Not null
Secquest	Varchar(50)	Not null
Secans	Varchar(50)	Not null

3. CONCLUSION

In this thesis, we have presented a system to filter undesired messages from OSN walls. The system exploits a ML soft classifier to enforce customizable content-dependent FRs. Moreover, the flexibility of the system in terms of filtering options is enhanced through the management of BLs. This work is the first step of a wider project. The early encouraging results we have obtained on the classification procedure prompt us to continue with other work that will aim to improve the quality of classification. In particular, future plans contemplate a deeper investigation on two interdependent tasks. The first concerns the extraction and/ or selection of contextual features that have been shown to have a high discriminative power. The second task involves the learning phase. Since the underlying domain is dynamically changing, the collection of pre classified data may not be representative in the longer term. The present batch learning strategy, based on the preliminary collection of the entire set of labeled data from experts, allowed an accurate experimental evaluation but needs to be evolved to include new operational requirements. In future work, we plan to address this problem by investigating the use of online learning paradigms able to include label feedbacks from users. Additionally, we plan to enhance our system with a more sophisticated approach to decide when a user should be inserted into a BL. The development of a GUI and a set of related tools to make easier BL and FR specification is also a direction we plan to investigate, since usability is a key requirement for such kind of applications. In particular, we aim at investigating a tool able to automatically recommend trust values for those contacts user does not personally known. We do believe that such a tool should suggest trust value based on users actions, behaviors, and reputation in OSN, which might imply to enhance OSN with audit mechanisms. However, the design of these audit-based tools is complicated by several issues, like the implications an audit system might have on users privacy and/or the limitations on what it is possible to audit in current OSNs. A preliminary work in this direction has been done in the context of trust values used for OSN access control purposes. However, we would like to remark that the system proposed in this paper represents just the core set of functionalities needed to provide a sophisticated tool for OSN message filtering. Even if we have complemented our system with an online assistant to set FR thresholds, the development of a complete system easily usable by average OSN users is a wide topic which is out of the scope of the current paper. As such, the developed Facebook application is to be meant as a proof-of-concepts of the system core functionalities, rather than a fully developed system..

REFERENCES

[1.] Renato M. Silva, Tulio C. Alberto, Tiago A. Almeid, Akebo Yamakami, “Expert Systems With Applications”, www.elsevier.com/locate/eswa, 2017.

[2.] Subramaniyaswamy V, Logesh R, Vijayakumar V, and Indragandhi V, “Data Mining for the Enterprise”, 2nd International Symposium on Big Data and Cloud Computing (ISBCC’15). Procedia Computer Science 50 (2015)

[3.] Kunal Jain, Sanjay Agrawal, U, “A Hybrid Approach for Spam Filtering”, 2014 5th International Conference- Confluence The Next Generation Information Technology Summit (Confluence).

[4.] N.Thilagavathi R.Taarika, “Content Based Filtering in Online Social Network”, New 2014 International Conference on Circuit, Power and Computing Technologies [ICCPCT].

[5.] Ms. Priyanka Salunkhe Mrs.Smita Bharne, Mrs.Puja PadiyaM, “Filtering Unwanted Messages from OSN Walls”, 2016 1st International Conference on Innovation and Challenges in Cyber Security (ICICCS 2016).