

Sentiment Analysis in Social Networking

Sudha P^{#1}, Abinaya K^{#2}, Janani S^{#3}, Mahalakshmi N^{#4}, SumaiyaAjrin S^{#5}

[#]Department of Computer Science and Engineering, SRM TRP Engineering College

¹sudhapanneerselvam20@gmail.com

²abinayakumar17@gmail.com

³janusaravanan06@gmail.com

⁴mahashruthi22@gmail.com

⁵sumaiyaajrin20@gmail.com

Abstract— Today's modern life is completely based on Internet. Now a day's people cannot imagine the life without Internet. Few years ago people share their views, ideas, information with each other using social networking sites. Such exchanges might include diverse sorts of substance such as text, image, audio and video data. One fundamental issue in today On-line Social Networks (OSNs) is to give users the ability to control the messages posted on their own private space to avoid that unwanted content is displayed. Hence Online Social Networks should be extremely secured and should protect the individual's privateness. The Online Social Network provides the security measures but it is only limited. While interacting the user can access the profile of other members which are involved in social sites and even share data such as images, text, videos etc. One critical issue in user wall is to give users the capability to control the messages posted on their own personal space in order to avoid unwanted content to be displayed on their wall. To overcome this problem, this is the proposed system allowing OSN users to have a direct control on the messages posted on their walls. This is achieved through a soft rule-based system, that allows users to personalize the filtering rules to be a matter-of-fact to their walls, and a Machine Learning based soft classifier automatically classify the messages in content-based filtering. It also includes the tagging based on the request analysis in online social network with real time indication.

Keywords— Short Text Classification, Blocking, Datamining, Social Networks, Rule-based system, Machine Learning

I. INTRODUCTION

Data mining is the estimating technique of discovering patterns in large data sets comprising methods at the intersection of machine learning, statistics, and database systems. It is a collaborative subfield of computer science. The overall goal of the data mining process is to extract information from a data set and convert it into an understandable structure for future use. Data mining is the examining step of the "knowledge discovery in databases (KDD)" process. Apart from the KDD[10], it involves data managing aspects, data pre-processing, model and inference considerations, interestingness metrics, complexity considerations, post-processing of discovered structures, visualization, and online updating.

In fig 1.1 describes the architecture for data mining. The absolute data mining work is the semi-automatic or automatic analysis of huge quantities of data to extract previously unknown, interesting patterns such as groups of data records (cluster analysis), exceptional records (anomaly detection), and dependencies (association rule mining, sequential pattern mining). This is normally used in database techniques such as spatial indices. These patterns can be seen as a kind of summary of the input data, and may be used in future analysis, machine learning and predictive analysis. As an instance, the data mining step might find out multiple groups in the data, which can be used to obtain more absolute prediction results with the help of decision support system. Neither the data collection, data preparation, nor result interpretation, the reporting plays a part in the data mining step where it belongs to the overall KDD process as an additional step.

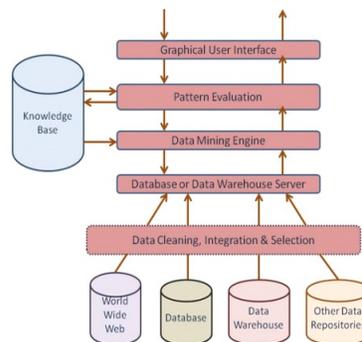


Fig. 1.1 Architecture for Data Mining

II. FACEBOOK

Facebook is a free social networking platform that encourage and facilitates interaction between friends, family and colleagues. It was founded by Mark Zuckerberg, along with fellow Harvard College students and roomie Eduardo Saverin, Andrew McCollum, Dustin Moskovitz and Chris Hughes. It is regarded as one of the Big Four technology companies along with Amazon, Apple, and Google.

The founders started with limiting the website's credentials to Harvard students and subsequently Columbia, Stanford, and Yale students. Credentials were eventually expanded to the remaining Ivy League schools, MIT, and higher education institutions in the Boston area. It moderately added support to students at various other universities, and eventually to high school students. Since 2006, anyone who affirms to be at least 13 years old has been granted to become a certified user of Facebook, though variations exist in this requirement, depending on local laws. The name comes from the face book records often given to American university students. It held its initial public offering (IPO) on February 2012, valuing the company at \$104 billion, the largest valuation to date for a newly listed public company. It began to sell its stock to the public three months later. It makes most of its revenue from advertisements that appear onscreen.

The service can be accessed from devices with Internet connectivity, such as personal computers, tablets and smart phones. After registering, users can create personalized profile revealing information about themselves.[9] Users can post text, photos, and multimedia of their own personal devices and share it with other users as "friends". Users can use various nested apps, and receive notifications of their friends' activities. Users may join common-interest groups.

It had more than 2.2 billion monthly active users as of January 2018. It receives prominent media coverage, including many arguments such as user privacy and psychological effects. The company has faced intense pressure over restrictions and over content that some users find objectionable.

III. PROPOSED METHOD

In content based filtering [4], each user is assumed to operate separately. As a result, this filtering system selects the information based on the correlation between the content of the information and the user preferences with similar choice. Documents processed in content-based filtering are mostly textual in nature. The activity of filtering can be modelled, indeed as a case of single label, binary classification, dividing incoming documents into relevant and non-relevant sections.

One fundamental issue in today On-line Social Networks (OSNs) is to give users the ability to control the messages posted on their own private space to avoid that unwanted content to be displayed. Till today OSNs provide little support to this requirement. To overcome the problem, in this paper, we propose a system allowing OSN users to have a direct control on the messages posted on their walls. This is fulfilled through a flexible rule-based system that allows users to customize the filtering paradigms to be applied on their walls.

Machine Learning [1] based soft classifiers automatically label messages in support of content-based filtering[4]. Machine learning (ML) is used as text categorization techniques to automatically assign each short text message within a set of categories based on its content. The major efforts in building a robust Short Text Classifier (STC) concentrate in the extraction and selection of a set characterizing and discriminating features. Here, a database of the categorized words is built and it is used to check out the indecent words. If the message consists of any vulgar words, then they will be sent to the Block list to filter out those words from the message. Finally, the message does not having any indecent words will be posted in the user's wall based on content-based filtering technique.

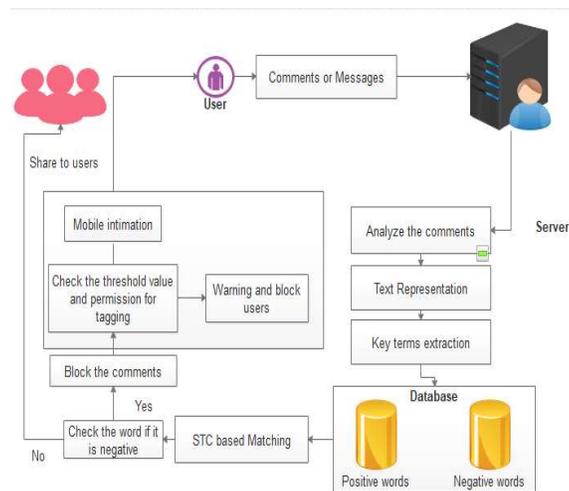


Fig 1.2 Architecture for Proposed Work

In Fig 1.2 describe the architecture for proposed work. In this architecture we can block the negative comments in social network page. The user can post the comments in page and send it to server. Server can implement text mining algorithm [2] used for eliminating the stop words. Finally extract the key terms and analyse the sentiment using short text classification algorithm. We can implement filtered rules based on threshold values and also block the friends who are continuously posting negative comments.

IV. IMPLEMENTATION AND RESULTS

A. Short Text Classification

From a ML point of view, we approach the task of short text categorization by defining a hierarchical two stage strategy

1) Hard Classification:

It is considered as a hard classification where short texts are labelled with crisp positive and negative text.

2) Soft Classification:

It acts on the crisp set of negative short texts and, for each of them, it simply yields estimated appropriateness or “gradual membership” for each of the conceived classes, without taking any hard judgments on any of them.

B. Blocking

The main implementation of our paper is to execute Automatic Blocking of comments and Creator, which will prevent comments from the undesired creators. If the Creators comments any negative words those comments will not be seen by other public instead it will remain in the user's profile and the User will be notified about this comments in Offline. If the creator exceeds the Threshold value which is given for negative comments he/she will be automatically blocked.

C. Tagging

In our paper we have done additional improvement to tagging that is the user will get offline notification about the tagging before the post is publicized to other users. Tagging is done only when the user accepts the tag request from other friends.

D. Offline Notification

There are some extreme scenarios where we totally reliant on notifications .In these cases, it is necessary to efficient in terms of energy or data spent on getting these notifications. In our paper we have provided offline notifications to the user about Comments and Tagging.

V. CONCLUSION AND FUTURE WORK

In this paper, we have analysed a system to filter unacceptable messages from OSN walls. The system accomplishes a ML soft classifier to enforce customizable content dependent FRS. The major efforts in building a robust short text classifier are concentrated in the extraction and selection of a set of characterizing and different features. This work is the first step of a broader 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 peculiarity of classification. In this system uses the ML soft classifier to remove the unwanted messages. Apart from classification facilities, the system provides a strong rule layer exploiting a flexible language to particularize Filtering Rules (FRs), by which users can state what contents, should not be exposed on their walls. In future we can further improve the classification process by adding some new features also. We have conferred extractive and abstractive caption generation models. As future work, we intend to exploit similar techniques to infer BL[10] rules and FRs[5]. In future, we can extend the framework to implement this system with various languages with improved accuracy [11]. Then using Semi-supervised approach [8] to predict the comments without trained databases.

REFERENCES

- [1] B. Wang, G. Chen, L. Fu, L. Song, and X. Wang, "Drimux: Dynamic rumor influence minimization with user experience in social networks," in Proc. 30th AAAI Int. Conf. Artif. Intell., Feb. 2016.
- [2] D. N. Yang, H. J. Hung, W. C. Lee, and W. Chen, "Maximizing acceptance probability for active friending in online social networks," in Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2013, pp. 713–721.
- [3] A. McCallum, A. Corrada-Emmanuel, and X. Wang, "Topic and role discovery in social networks," in Proc. 19th Int. Joint Conf. Artif. Intell., 2005, pp. 786–791.
- [4] L. Fu, W. Huang, X. Gan, F. Yang, and X. Wang, "Capacity of wireless networks with social characteristics," IEEE Trans. Wireless Commun., vol. 15, pp. 1505–1516, Feb. 2016.
- [5] A. Montanari and A. Saberi, "The spread of innovations in social networks," in Proc. National Academy of Sciences of the United States of America PNAS, Aug. 2010, pp. 20 196–20 201.
- [6] X. Rong and Q. Mei, "Diffusion of innovations revisited: From social network to innovation network," in Proc. 22Nd ACM Int. Conf. Inf. Knowl. Manag., 2013, pp. 499–508.
- [7] C. Budak, D. Agrawal, and A. E. Abbadi, "Limiting the spread of misinformation in social networks," in Proc. 20th Int. Conf. World Wide Web, 2011, pp. 665–674.
- [8] A. Bessi, F. Petroni, M. Del Vicario, F. Zollo, A. Anagnostopoulos, A. Scala, G. Caldarelli, and W. Quattrociocchi, "Viral misinformation: The role of homophily and polarization," in Proc. 24th Int. Conf. World Wide Web, 2015, pp. 355–356.
- [9] E. Serrano, C. A. Iglesias, and M. Garijo, "A novel agent-based rumor spreading model in twitter," in Proc. 24th Int. Conf. World Wide Web, 2015, pp. 811–814.
- [10] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in Proc. 9th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2003, pp. 1175–1180.
- [11] Sang-Hun Cho and Hang-Bong Kang, "Text Sentiment Classification for SNS-based Marketing Using Domain Sentiment Dictionary" in Proc. The EMNLP, pp.79-86, 2002.
- [12] K. Mouthami, K. Nirmaladevi, Dr. V. Murali Bhaskaran, "Sentiment Analysis and Classification Based On Textual Reviews" in proc. 3rd IEEE international conference on data mining (ICDM), USA, pp.427-434.