

A Survey To Detect Undesirable Messages In Online Social Network

Manoj Prabhakar J, Dr.P Ezhilarasu , Satheesh Kumar D , Sowmya R

Abstract— In this paper, we discuss about various methods to remove undesirable messages in social media. A lots of papers are taken for the survey. Its advantages and disadvantages are discussed.

Keywords— Online social network, Filtering methods, Text classification

I. INTRODUCTION

Online Social Network is one of the popular interactive medium to share our views, communicate with peoples, meet new friends. The social network perspective provides a set of methods for analyzing the structure of whole social entities as well as a variety of theories explaining the patterns observed in these structures.[1] It also includes free text, image audio and video data. These sort of data can be shared to every users all around the world. According to face book user shares more than 10 billion piece of contents such as web links, news stories, blog posts, notes, photo albums, videos are shared every month. The main aim of these proposals is to prevent the users from unwanted messages and vulgar images in online social network. In Online Social Network, Information filtering is used to automatically control the messages posted on their walls. In this paper short text classification technique is used for text classification it classifies the given texts as neural and non-neural words. It calculates as follows

- Neural words
Set of un important words
- Vulger words
Set of unwanted texts
- Capital words
Total number of words in capital letter
- Punctuations

It calculates the amount of punctuations used in the comments

- Special symbols

Exclamation marks, question marks and other symbols that are used .

The main process that is carried out in non-neural technique is categorize the vulgar or bad comments that are posted on the user walls. We introduce Machine Learning (ML) for text categorization techniques to automatically assign with each short text message a set of categories based on its content. In this paper we propose black list management so that the user can prevent from unwanted posts.

EXAMPLE OSN like Facebook, Twitter, Douban and other social medias makes way for user to insert messages in their walls. It also cannot prevent the unwanted posts posted on their walls. For this purpose these methods are undertaken here. Its overall performance are being analysed.

II. VARIOUS MESSAGE FILTERING METHODS

TABLE 1
Analysis of message filtering methods

S.NO	TITLE	DESCRIPTION
1	Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extension[2]	In this method an overview of three main Categories content-based, collaborative, and hybrid recommendation approaches are used to filter messages
2	A Machine Learning Approach to WebPage Filtering Using Content and Structure Analysis[3]	In this paper we propose a machine-learning-based approach that combines Web content analysis and Web structure analysis. We represent each Web page by a set of content-based and link-based features, which can be used as the input for various machine learning algorithms
3	Content-Based Book Recommending Using Learning for Text Categorization[4]	In this paper we describe a content-based book recommending system that utilizes information extraction and a machine-learning algorithm for text categorization

Manoj Prabhakar J ,PG scholar, Department of Computer Science and Engineering Hindusthan College of Engineering and Technology, Coimbatore, Tamilnadu, India manojprabhakarj92@gmail.com

Dr .P Ezhilarasu ,Associate Professor, Department of Computer Science and Engineering Hindusthan College of Engineering and Technology, Coimbatore, Tamilnadu, India prof.p.ezhilarasu@gmail.com

Satheesh Kumar D, Associate Professor, Department of Computer Science and Engineering Hindusthan College of Engineering and Technology, Coimbatore, Tamilnadu, India dsatheeshme@gmail.com

Sowmya R ,PG scholar, Department of Computer Science and Engineering Hindusthan College of Engineering and Technology, Coimbatore, Tamilnadu, India sowmyacsebe@gmail.com

4	Machine Learning in Automated Text Categorization[5]	In this method automated text categorization are taken to find out neural and non-neural words
5	Content-Based Filtering in On-Line Social Networks[6]	This system allows OSN users to have a direct control on the messages posted on their walls. This is achieved through a flexible rule-based system
6	Information Filtering and Information Retrieval[7]	This system deals primarily with textual information, but they may also entail images, voice, video or other data types that are part of multimedia information systems
7	Personalized Information Delivery: An Analysis of Information Filtering Methods[8]	In this Two information retrieval methods were tested to make the predictions, one using standard keyword matching, and the other using Latent Semantic Indexing (LSI). All four methods effectively selected relevant abstracts
8	Scisor: Extracting Information from On-Line News[9]	In this paper, Drawing on artificial intelligence techniques, and applying them to financial news items, this powerful tool illustrates some of the future benefits of natural language analysis through a combination of bottom-up and top-down processing
9	A Rule-Based Message Filtering System[10]	This paper describes the ISCREEN prototype system for screening text messages. ISCREEN includes a high-level interface for users to define rules, a component that screens text messages
10	Tcs: A Shell for Content-Based Text Categorization[11]	TCS consists of the TCS run-time system and a rule base. The rule base defines what categories the application can assign to texts and contains rules that make the categorization decisions for particular text

and unwanted messages posted on their walls. Due to this many users deactivate their user accounts in online social network. In order to overcome this problems these methods are introduced. The result of these methods are analyzed and it shows that A Rule Based Message Filtering System is better method to detect the unwanted messages.

REFERENCES

- [1] https://en.wikipedia.org/wiki/Social_network
- [2] A. Adomavicius and G. Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions," IEEE Trans. Knowledge and Data Eng., vol. 17 no. 6, pp. 734-749, June 2005
- [3] M. Chau and H. Chen, "A Machine Learning Approach to Web Page Filtering Using Content and Structure Analysis," Decision Support Systems, vol. 44, no. 2, pp. 482-494, 2008.
- [4] R.J. Mooney and L. Roy, "Content-Based Book Recommending Using Learning for Text Categorization," Proc. Fifth ACM Conf. Digital Libraries, pp. 195-204, 2000.
- [5] F. Sebastiani, "Machine Learning in Automated Text Categorization," ACM Computing Surveys, vol. 34, no. 1, pp. 1-47, 2002.
- [6] M. Vanetti, E. Binaghi, B. Carminati, M. Carullo, and E. Ferrari, "Content-Based Filtering in On-Line Social Networks," Proc. ECML/PKDD Workshop Privacy and Security Issues in Data Mining and Machine Learning (PSDML '10), 2010
- [7] N.J. Belkin and W.B. Croft, "Information Filtering and Information Retrieval: Two Sides of the Same Coin?" Comm. ACM, vol. 35, no. 12, pp. 29-38, 1992
- [8] P.W. Foltz and S.T. Dumais, "Personalized Information Delivery: An Analysis of Information Filtering Methods," Comm. ACM, vol. 35, no. 12, pp. 51-60, 1992
- [9] P.S. Jacobs and L.F. Rau, "Scisor: Extracting Information from On-Line News," Comm. ACM, vol. 33, no. 11, pp. 88-97, 1990.
- [10] S. Pollock, "A Rule-Based Message Filtering System," ACM Trans. Office Information Systems, vol. 6, no. 3, pp. 232-254, 1988.
- [11] P.J. Hayes, P.M. Andersen, I.B. Nirenburg, and L.M. Schmandt, "Tcs: A Shell for Content-Based Text Categorization," Proc. Sixth IEEE Conf. Artificial Intelligence Applications (CAIA '90), pp. 320-326, 1990.

III. CONCLUSION

In this paper we studied about various message filtering methods to prevent the users from interacting with strangers