

INFORMATION FILTERING ON SOCIAL MEDIA SITES TO AVOID COMMUNAL AND POLITICAL ISSUES

Vishal Kaushik, UPES, Dehradun, India
(vkaushik@ddn.upes.ac.in)

Akshay Toshniwal, UPES, Dehradun, India
Durgesh Sharma, Paragon Science Academy, USA
(dsharma@paragonscience.org)

ABSTRACT

Social Media is an emerging form of communication that has become very popular in recent years. It allows users to publish updates as short text messages and broadcast it to followers. Currently Twitter, Facebook, LinkedIn, Yahoo Meme and many more websites are excellent instance of socializing Medias. Through them, not only personal communication, but also, a transfer and sharing of political and general events is exhibited. Unfortunately, such liberty rendering service also has its drawbacks. Today's social media websites do not provide an ability to control the contents of the message; any unwanted contents can easily be shared by millions of users, just a click away. Therefore, security in sharing of such illegitimate messages is an important issue with such websites. Facebook and Twitter are some of the biggest social media service used all over the world to broadcast views, opinions about related topics. Due to the lack of classification system that winnows unwanted messages from appropriate ones, the user receives all kinds of updates and posts which can lead to various political and communal issues. In this paper an information filtering for twitter is introduced. We hereby propose an online filtering system, to filter messages that can hurt the religious sentiment of the users and often time lead to political issues [1].

Keywords: Social Media analytics, Communal harmony, Information Mining, Content Categorisation, Content – scenario analysis, Sentiment Analysis.

INTRODUCTION

Web 2.0 was often defined as a second generation in the development of World Wide Web (WWW), forming a combination of concepts, technologies and additional features that focused on sharing of user-generated content and social networking. Social media is on its zenith and its usage has increased phenomenally; a primary cause could be the huge and rapid advancement if computer and information technology. People are using social media on a daily basis, to share their opinions with each other on a wide variety of subjects, products and services. Sometimes to provide reviews, feedback, and posting comments on various issues, making them one of the most popular interacting medium. The relation between the web authors and the web users has changed greatly by providing users, the ability to create their own contents. Facebook, LinkedIn, Twitter and many more social media websites, constitute infrastructures created by user based activities, extracted by these sites. One of the main features of micro-blogging and Social Network Sites is that users do not just produce contents, rather they can be a part of conversations with other users through commenting, liking and sharing their posts. Within this social media conversation, it is necessary to understand the informative content as well as the context of the information being posted. In this paper, we will filter out the messages in Twitter, based on the words contained in it; words that could be harmful to the communal, political or social values of the person. At present, there is no filtering mechanism which blocks unwanted and/or illegitimate messages posted by the user. Thus, for a better security, avoidance of false rumors, and preventing communal and political issues, a filtering mechanism is proposed, which will block the unwanted messages from being posted [1][2].

ISSUES

Issues in the social media are rising day by day, due to posting of unwanted and illegitimate messages and opinions on the walls of the friends or followers. This has led to many social and political issues and also has harmed the society. Some of the legal issues that can be caused by just posting of someone's messages are as follows:

1. Disclosure of confidential information.
2. Defamation issues.
3. Creation of Fake IDs.
4. Unauthorized use of copyright-protected works.

The major usage pattern of social networking sites/ blogs in India is categorised/ analysed. Based on the analysis done, it has been found that the maximum number of religious or communal issues are arising due to the Social Media. No one is to blame about this, but since there is no filtering mechanism over social media, users are allowed to post irrelevant as well as unwanted messages [3].

RESEARCH METHODOLOGY

Social media has been increasing day by day and has taken a new leap in the field of spreading of information and has become a part of media. Before broadcasting it on the medium such as radio, television, etc. social media has now been a part of media. It helps in broadcasting the latest updates related to every field be it, political, social, technology and every aspect of the society. Worldwide news is broadcasted within seconds.

The present scenario in social media is that the information that are posted on the social media are not filtered or checked for any unwanted texts or irreverent data. This leads to various social and political issues, which can be harmful to the country and the sentiments of the users as well. To avoid this a filtering mechanism is proposed. The approach and the algorithm for the system is given below:

Naïve Bayesian Algorithm

Naïve Bayes algorithm represents a supervised learning method for classification. It was introduced for text retrieval and categorisation. Assumes an underlying probabilistic model and it allows us to capture uncertainty about the model in a principled way by determining probabilities of the outcomes. It can solve diagnostic and predictive problems. The algorithm will be applied at the following two phases:

1. Fetching the tweets or the messages posted on the social media websites and analysing them.
2. For analytical purposes and to categorise the text based on the words present in the message broadcasted or posted by the user.

The Naïve Bayesian algorithm mathematically can be expressed as shown below: Bayes theorem provides a way of calculating the posterior probability, $P(c|x)$, from $P(c)$, $P(x)$, and $P(x/c)$. Naive Bayes classifier assume that the effect of the value of a predictor (x) on a given class (c) is independent of the values of other predictors. This assumption is called class conditional independence.

$$P(c | x) = \frac{P(x | c) P(c)}{P(x)}$$

Likelihood
Class Prior Probability

Posterior Probability
Predictor Prior Probability

$$P(c | \mathbf{X}) = P(x_1 | c) \times P(x_2 | c) \times \dots \times P(x_n | c) \times P(c)$$

PROPOSED SYSTEM

The proposed system is shown in the figure given below:

It consists of the following steps:

1. **Data Access:** The messages that are posted on the walls of the user will be retrieved by using the following API which is suitable for that social media website.
2. **Data Cleaning:** Data cleaning will be done using text mining and the special symbols, emoticons, etc. will be filtered out for better results. The algorithm used for text mining can be Naïve Bayes.
3. **Data Analysis:** The data i.e. the text which is filtered using text mining will be analysed for suitable words and statements which can prove harmful to the society and the community. These words can be of the following types and category [4][5]:

Category	Words
Terrorism	Bomb Blasts, Suspect, etc.
Political	Defamation, Communalism, Riots, etc.
Vulgar	Sex, Fuck, etc.

The text categorization (data analysis) will be done using Naïve Bayesian Classifier, which is used to classify the texts into different and appropriate category. It provides 75%-85% accuracy and thus Naïve Bayesian Classifier is used.

Lastly after analyzing the complete text or the message, the message will have certain set of words which can be harmful to the nation's security and can cause communal and political issues as well. Thus to prevent such issues and to enhance the security measures of the social media in posting messages, the system will block those messages which can cause any harmful effect to the society and the country.

Our system will also block those users which are continuously posting such messages which can lead to any of the issues mentioned above. Based on the content of the message the user has posted, and also on the frequency of the user to post such message, the user will be block for a specific amount of period [3][4].

CONCLUSION AND FUTURE WORK

The system implemented will thus provide and enhance the security over social media in preventing the unwanted and illegitimate messages. This will lead to avoidance of causing any political and communal issue and will also help the cyber law to detect any unwanted activity by users in posting such kind of messages. The system can be further upgraded by retrieving the images also from the message. Now a days the messages contains images, which depicts a lot of information. Thus the future work will be classifying the images which will help in enhancing the security measures at a greater extent [4][5].

REFERENCES

- [1] BarisGuc, A. (2010). Information Filtering on Micro-blogging Services, Swiss Federal Institute of Technology Zurich.
- [2] Rakhi Bhardwaj, Vikaram Kale, Prasad Morye, ManojDhaygude, SagarBadhe, Unwanted Message Filtering from OSN user walls and Implementation of Blacklists, International Journal of Engineering and Computer Science, ISSN:2319-7242.
- [3] Matteo Magnani, DaniloMontesi, Luca Rossi, Conversation Retrieval for Microblogging Sites.

- [4] Eman M.G. Younis, F. (2015). Sentiment Analysis and Text Mining for Social Media Microblogs using Open Source Tools: An Empirical Study, *International Journal of Computer Applications* (0975-8887).
- [5] HimanshuSuyal, M. (2014). RB Patel, Improved Information Filtering and Feature Dimensionality Reduction using Semantic based Feature Dataset for Text Classification: In Context to Social Network, *International Journal of Computer Applications*, (0975-8887).
- [6] Anonymous, Naïve Bayesian Classification Algorithm.