Social Network Based Disaster Analysis Using Naive Bayesian Classification Algorithm

[1]T. Rajasekaran[*], [2]N. Raveena, [3]S. Revathi, [4]J. Sruthi
[1] Assistant Professor, [2, 3, 4]UG Scholar
[1, 2, 3, 4] Department of Computer Science and Engineering, KPR Institute of Engineering and Technology, Arasur, Coimbatore, Tamilnadu, India

Abstract--Nowadays each and every one having their account in various social networks. Social network is used to spread the massive information during emergency situation. In this paper we are going to create a trusted network for disaster, So that the information shared considered as trusted and fake information's can be avoided. The ultimate aim of this paper is to warn and create awareness among the public in earlier, so that they can be able to rescue themselves from various disasters by saving their lives and property. The information about each and every disasters is posted in social network by scientists or research people from various research centers. By using Naive Bayesian Classification algorithm we can be able to analyze whether the information is trusted or fake, and the dictionary is maintained for analyzing trusted or fake posts, and it is used to find the accuracy for each posts. In this method only the trusted posts is taken for consideration and a warning will be sent the people in particular area, through social network.

Keywords--Social network, Disaster, Naïve Bayesian Classification algorithm, Research center, Post, Awareness.

I. INTRODUCTION

Social network is mainly used to stay up-to-date with news and current events, to share our opinion's, to find some funny news and entertainment, to fill up some spare time, etc. Social network plays an important role in spreading the information's quickly during emergency situations [8]. Almost 80% to 90%of people have their account in various social networks like face book, twitter, linked in, you tube [4]. In this paper we are using social network to post the disaster information to create awareness to the public prior. We need to create the trusted social network for reducing the fake information's. The disaster information will be sent by research centers to the public who belong to that prone area, this is mainly used to avoid rumors [3]. After discussing with various scientists only the information will be posted in social networks. The various researchers, scientists, scholars will analyze the disaster, and information will be posted in social network. The information can be viewed by public if they belong to that particular core area. Text classification is used to classify each and every post and it is mainly used for sentimental analysis[2]. Here Naïve Bayesian Classification algorithm is used for text classification, sentimental analysis, and for spam filtering [1]. It is used to find the positive and negative posts.

II. RELATED WORK

The existing system will post the disaster information in the social network so that each and every people can be able to view and forward the post. Sometimes fake information also will be shared. To overcome the problem we are using the -proposed system. In proposed system we are going to create the trusted network, so that the disaster related information will be shared to the person who belongs to that prone area. Only research center posts will be allowed in the trusted network. Machine learning algorithm is used to classify each and every posts by using text classification. The Classifier algorithm is used to train and test the data to find the trusted, fake posts and the accuracy of each and every posts [8]. The dictionary is also maintained to avoid rumors. Each and every post from the research center will be analyzed, if the keywords mismatched means the information will not be post successfully.

III. DATASETS

Fig.1 Data sets maintained in database

In this project we are considering the unstructured data in the form of text format. The example data sets are "Tsunami may occur by today evening from 12.PM to 1.00 PM". These type of data sets will be stored in the database and used for analyzing.

IV. PROPOSED WORK

A. *Social Network Creation*

Social network is a web application that has been created for each and every individual to create their profile to share their information's, etc. Common web applications include web mail, online chatting, image sharing, online retail sales, online auctions and many other functions. Here we have created a trusted network for posting disaster information to the people who belongs to that prone area. This can be able to reduce the rumors.

Fig.2 User Login Form

B. *Trustworthy Monitoring*

This is the core module in our project. Here each and every post would be considered as trusted, because we are considering only the research center post. Based on the repetitive post about the same information from various research centers will be taken into considerations and it will be sent to the people who belong to that core area in prior. Dictionary is also maintained to analyze the trusted information posted from the research centers. For example some scholars or research people might post that "We need to pray for those who died in various disasters", this is not the trusted post to alert the people [7]. This type of information can be able to create rumors, so dictionary is maintained to restrict these type of information's.

Fig.3 Notification sent to particular area people

C. *Naive Bayesian Classification*

The classification algorithm is used to train the number of posts. Based on the training data each and every data will be tested and classified to find the positive and negative words in it.The Naïve Bayesian Classification **is** a technique based on Bayes Theorem

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)}$$

The Naïve Bayesian Classification algorithm is mainly used for prediction, text classification, sentimental analysis, spam filtering, etc [10]. Here we are going to find the polarity of each and every sentences posted by research center in the social network. The polarity of each and every sentence can be able to identify in this algorithm [5]. In Naïve Bayesian Classification algorithm we can be able to predict the problem and accuracy of certain problem can be found easily and effectively.

Fig.4 Finding Accuracy for each posts

V. RESULTS AND DISCUSSION

We have created a trusted network in ASP.NET and the disaster information will be posted by various research centers to the people who belong to that core area. The data will be stored in database, the data will be analyzed and classified by using text classification and Naïve Bayesian Classification. This will be done in Linux to find the polarity of each sentences. We have to integrate the Windows and Linux to predict the output in the form of graph by using Matplotlib.
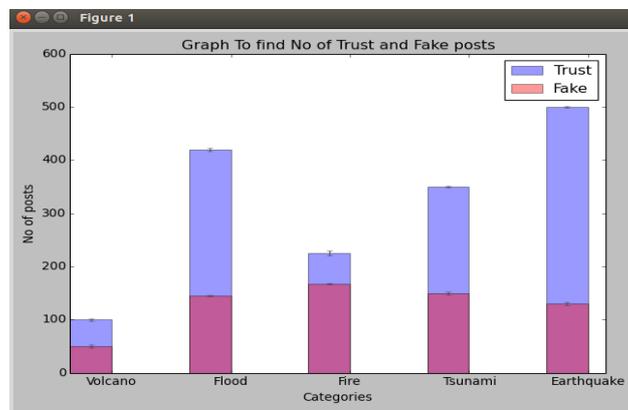

Fig.5 Graph for No of trust and fake posts

VI. CONCLUSION

In this paper we have created a trusted network and only the disaster related information will be shared to the registered user who belong to that prone area. After analyzing the posts we have found the sentimental analysis by using text classification method and the accuracy is also measured by machine learning algorithm. Based on the number of Affirm (positive posts) and Flaw(negative posts) the information will be given as alert to the registered user in prior, so that the people can be able to rescue themselves quickly.

REFERENCES

[1]   Bharath Sriram, David Fuhry, Engin Demir, Hakan Ferhatosmanoglu, Murat Demirbas, "Short Text Classification in Twitter to Improve Information Filtering", *SIGIR'10 ACM, july 19-23, 2010.*

[2]   Chu-Ren Huang , Shoushan Li, Sophia Yat Mei Lee, Zhongqing Wang, "Sentiment Classification with Polarity Shifting Detection" in Data Mining, 2003. ICDM2008.Third IEEE International Conference on. IEEE, 2008, pp. 427–434.

[3]   Dipali D. Vidhate, Ajay. P. Thakare, "A System Approach to Avoid Unwanted Messages from User Walls", *International Journal of Science and Research(IJSR) Vol 3 Issue 2, Feb 2014.*

[4]   [Karimi et al., 2013] S. Karimi, J. Yin, and C. Paris. Classifying microblogs for disasters. In Proceedings of the 2013 Australasian Document Computing Symposium, pages 26–33, Brisbane, Australia, 2013.

[5]   Pang, B. and Lee, L. (2004). A sentimental education: Sentiment analysis using subjectivity analysis using subjectivity summarization based on minimum cuts.

[6]   Sharvil Shah, K Kumar, Ra. K. Saravanaguru "Sentimental Analysis of Twitter Data Using Classifier Algorithms".

[7]   Vanithasri, K. and Tamilmani.G "Natural Disaster Recommendation Speedy Emergency Alert system using Twitter for social networks".

[8]   [Yin et al., 2012] J. Yin, A. Lampert, M. Cameron, B. Robinson, and R. Power. Using social media to enhance emergency situation awareness. IEEE Intelligent Systems, 27(6):52–59, 2012.

[9]   Zielinski, A. and Bügel, U. (2012) Multilingual Analysis of Twitter News in Support of Mass Emergency Events, *Ijiscram*.

[10]   [Yin, 2013] J. Yin. Clustering microtext streams for event identification. In Proceedings of the 6th International Joint Conference on Natural Language Processing, pages 719–725, Nagoya, Japan, 2013.