Disaster Reporting and Alert System Using Tweets in Social Media

P.Tamije Selvy, V. Suriya Prakash, S. Shriram, N. Vimalesh, M. Anitha

*Abstract***:** *Social media is one of the powerful micro blogging platforms which is used to express the wide range of thoughts by an individual. Since this type of vast data is pretty much helpful to bring various applications. The pre-existed model does not comprise dynamic data generated by the user, but this model is programmed in a way to extract the dynamic data that is generated by various users. This proposed model investigates the real-time interaction of events such as earthquakes, tsunami, etc., in social media and proposes an algorithm to monitor hashtags and to report disaster. The model reports a disaster and gives alert to the users residing in the disaster location using the words in the posts relating the disaster event, number of words, and their context. The model considers each hash tags as an input and applies semantic analysis, which is widely used for estimation of consequences. Because of the numerous disasters and large number of Social media users throughout the country, this proposed system can report disaster more accurately by monitoring hashtags. The model reports natural disasters give alert to the users and also represent it visually.*

*Index Terms***:** *micro blogging, user tweets, twitter.*

I. INTRODUCTION

Social Media, a popular micro blogging service, has received much attention recently. The online social network is used by millions of people around the world to remain socially connected to their friends, family members, and co-workers through their computers and mobile phones. A status update message, is often used as a message to friends and colleagues. A user can follow other users; that user's followers can read her posts on a regular basis. A user who is being followed by another user need not necessarily reciprocate by following them back, which renders the links of the network as directed. Social media users have increased rapidly. For example, the number of registered Twitter users exceeded 100 million in April 2010. The service is still adding about 300,000 users per day.1 Currently, 190 million users use Twitter per month, generating 65 million tweets per day Many researchers have published their studies of Twitter to date, especially during the past year. Most studies can be classified into one of three groups: first, some researchers

have sought to analyze the network structure of Twitter. Second, some researchers have specifically examined characteristics of Twitter as a social medium Third, some researchers and developers have tried to create new applications using Twitter .Twitter is categorized as a micro blogging service. Micro blogging is a form of blogging that enables users to send brief text updates or micromedia such as photographs or audio clips. Micro blogging services other than Twitter include Tumblr, Plurk, Jaiku, identi.ca, and others. 3. in this study, which is based on the real-time nature of one social networking service, is applicable to other micro blogging services, but here specifically examine Twitter in this study because of its popularity and data volume . In another instance, when an airplane crash-landed on the Hudson River in New York, the first reports were published through Twitter and tumblr. The several messages that are share daily can be used for many good and as well as bad issues. Twitter is used for the transfer of messages on different terrorist attacks by the ISIS jihadis [1]. Hundreds of ISIS-supporting accounts sent tweets with location metadata embedded [2]. This model which also uses a micro blogging application has similar features like twitter. The dynamic data is generated by various users for a particular event, this generated data is pretty much helpful to generate a report for the particular event. This model highly concentrates on disaster events. The data generated by the user in the time of disaster is pretty much valid to know the immediate cause of the disaster. This model extracts the data generated by various user and which examine those data with a algorithm which coordinates the valid data regarding the particular disaster, this valid data are used to find the aftermath or consequences of the occurred disaster which also programmed in a way to sent the alert message to the particular user who belongs to the disaster area. The results are however produced as a graphical representation.

II. RELATED WORKS

Chatfield A, has made a complete twitter data analysis by using the social network analysis where the datasets collected from is completely searched for any types of messages that includes terrorist messages. They found a strong evidence for Sham witness-intermediated multi-sided Twitter networks of international mass media, regional Arabic mass media, IS fighters, and IS sympathizers, supporting the framework's utility [3]. Johoo Kim and Makarand Hastak also used the social network analysis for analyzing the twitter data after a

 Dr.P.Tamije Selvy, Department of Computer Science and Engineering, Sri Krishna College of Technology, Coimbatore, India.

 V.Suriya Prakash, Department of Computer Science and Engineering, Sri Krishna College of Technology, Coimbatore, India.

 S.Shriram, Department of Computer Science and Engineering, Sri Krishna College of Technology, Coimbatore, India.

 N.Vimalesh, Department of Computer Science and Engineering, Sri Krishna College of Technology, Coimbatore, India.

 M.Anitha, Department of Computer Science and Engineering, Sri Krishna College of Technology, Coimbatore, India.

disaster and to generate a report for the same where the main concentration goes on the graphical representation of the twitter data and they have also came to conclusion by doing a comparative study between facebook and twitter and giving the rate of posts that are shared in both the social medias and gave a report in a graphical way, they have also mentioned about the Louisiana floods and proved their point with the same as an example[4], however they produced a report after a disaster has occurred but no information was produced at the time of disaster, this paper gives the alert to the users at the time of disaster and also gives a graphical report.

Wu He, Shenghua Zha and Ling Li proposed a competitive analysis and text mining on social media. They proposed a case study in the pizza industry. The three pizza industries Pizza hut, Dominos pizza and Papa john's pizza are compared using the posts that are posted in the twitter. Here the data are taken and the comparative study of the number of tweets with respect to the pizza industries. Here text mining algorithm is used and the entire tweets about the pizzas are scanned and then the tweets that match with these pizza industries are extracted and then the comparative study is reported [5].

III. PROPOSED SYSTEM

In this model the posts that are posted by the users are the datasets which are then extracted from the database using an algorithm and then the accuracy of the data is checked using a centralized database and then an alert message is sent to the users residing in the disaster location. The proposed model consists of the following modules as depicted in Fig 1.
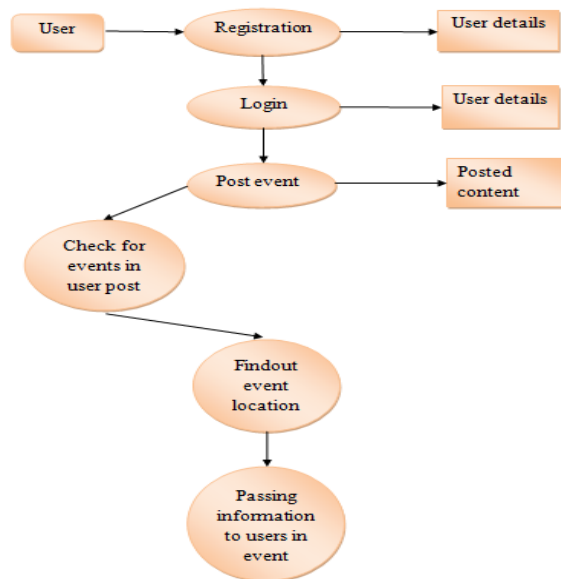

Fig 1.Basic block diagram

A. Social media creation

A social media is a one where a large number of users share their posts on a public cause or their personal things which can be viewed by their friends in the social media. These types of social media have become popular in recent times and there are nearly 2.77 billion social media users all over the world. Hence the data in the social media can be used for many good causes. Nowadays the information that are posted in the social reaches more number of people in shorter time than the information that are given by media, hence some of the

valuable information that are posted in the social media can be used for alerting the people about a disaster and it can also be used for generating a report for the consequences of a disaster. Due to some security reasons in the recent times the social media like 'twitter' are not giving out their data since it may affect the privacy of their users and hence for the processing of this paper a social media model has to be created which acts like a social media where the datasets are collected. The social media created is done completely using java, to be more specific java servlet programming is used to design the user interface in the social media like the registration page, login page, etc., [6]. Java servlets are the programs that run on the server side which return a dynamic or customized response. HTTP protocol is used to run the java servlets. The servlets are the ones that handle complex requests that are obtained from the web server. Servlet is comparatively faster than CGI because in servlet for every new request a new process is not created. Hence java servlet programming is more efficient and it is used to create the social media model. The newsfeed of the model is represented in the Fig 2. For the backend database connectivity JDBC connectivity is used. JDBC comes under the Java Standard Edition. The connection of database is established using the JDBC drivers and the MySQL queries are used for creating the structure of the database where the details of the user are stored and also the posted content is stored [7]. The structure of the database is represented in the Table I.
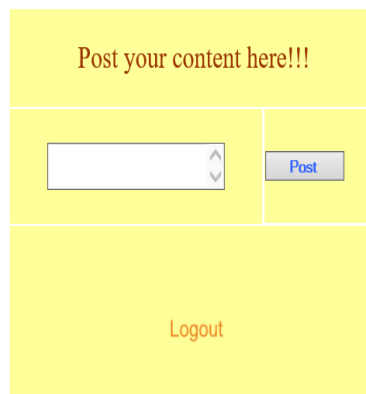
Fig 2.Social media newsfeed

Table I.
Database Structure Of User Details

Field Name	Data type	Size
Name	Varchar	50
Username	Varchar	50
Password	Varchar	50
Profession	Varchar	50
Address	Varchar	50
City	Varchar	50
Mobile	Number	10
Country	Varchar	50
Email id	Varchar	50

B. Data extraction

The datasets for this model are the posts that are posted by the user which can be extracted from database using semantic analysis. Semantic analysis is nothing but the given statement comes under the certain rules that are predefined in the same way a certain keywords that are related with the disaster like earthquake,flood,etc,. are given in the database using which the posts that match with the keywords are extracted from the database. This extraction process is done using the naïve bayes algorithm. Naïve bayes algorithm uses the bayes theorem for the classification of the datasets. This algorithm was introduced in 1960. This bayes theorem shown in (1) is based on the concept of probability and hence it gives a probabilistic result which has a high probability [8].

$$p(C_k|x) = \frac{p(C_k)p(x|C_k)}{p(x)} \tag{1}$$

x - problem instance
C_k – k possible outcomes

the conditional distribution p(x) in (2) over the class variable C [9] is,

$$p(x) = \sum_k p(C_k)p(x|C_k) \tag{2}$$

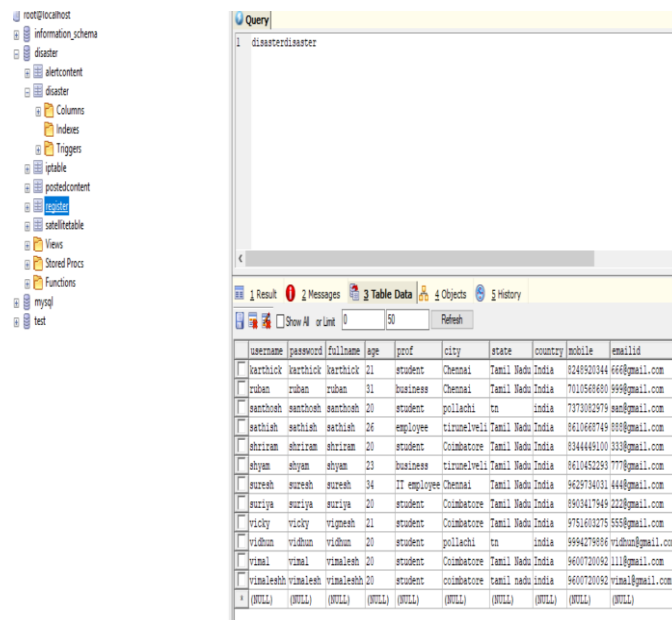p(x) – Scaling factor dependant on x

Fig 3. Database structure of user details

C. Alert system

When a disaster occurs in a specific location an alert to the people residing in the places should be given. There are many devices which can predict the disaster before it occurs but even that devices can't predict the disasters accurately and also even if the devices predict the disaster the locations that are very close to the disaster location is alone alerted by the government this doesn't cover the location that are a little far from the disaster location which may als get affected. For example if a tsunami occurs only the coastal areas are alone alerted by the government and hence by using a social media we can take the alert system to the next level by giving the alert to the users with the help of information from the messages posted in the social media. This can be successful because a information in a social media spreads more faster than information on media. All the messages that are posted in a social media cannot be taken as a true one there may be some fake news also spreading in the social media and hence the correctness of the information has to be verified. When a disaster occurs the information of the disaster is immediately entered in a centralized database which is maintained by the government hence the information that is gained from the social media has to be cross verified with this centralized database by matching the time and date of the disaster, if this details match with one another then an alert is sent to the users residing in the disaster location in a form of a text message. This can be achieved in java using a function way2sms which matches the IP address of the system where the application is hosted and the IP address of the users mobile phone[10]. The mobile number of the user is given at the time of registration and using this mobile number the system generated sms is sent to the users mobile. This gives the alert not only to the users those who are very near to the disaster location but also to the users who are a little far from the location. This alert system will reach a lot of people compared to the information that is transferred by the media.

Table II
Structure Of Central Database

Field Name	Data type	Size
Area	Varchar	50
Disaster name	Varchar	200

D. Data Visualization

Data visualization is nothing but representing the given datasets in a graphical representation which would be easy to understand. Here the datasets are represented in a graphical representation with respect to the location. The datasets that is the messages that are posted by the users are stored in the database and these data are used to represent the data in a chart representation. The chart is generated according to the number of tweets that are tweeted in a particular location about a disaster in a location. These data are represented in different types of charts. For example, a if a pie chart is generated a the number of tweets in a location is represented with respect to the location from where the tweet is posted. A pie chart usually gives the exact amount and the representation of the data given where a circle is divided into different parts and the report is generated which is easily understood. Here for the chart generation JFreeChart function in java is used. JFreeChart is a free java chart library which contains different types of charts which gives a easy way to display the given datasets in a chart representation [11]. This java library which contains several charts gives the output by just calling the function with the name of the chart type in which the output has to be displayed. Hence the Jfreechart library can be used to generate the chart with the datasets that is given by the previous module that is the datasets that have been verified with the centralized database.
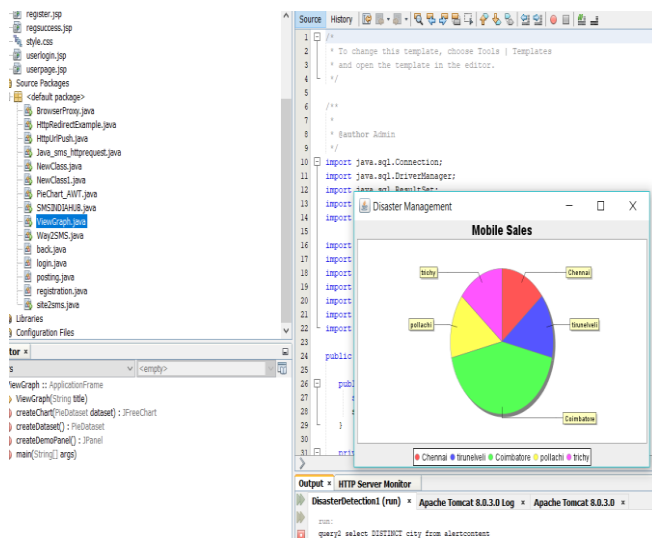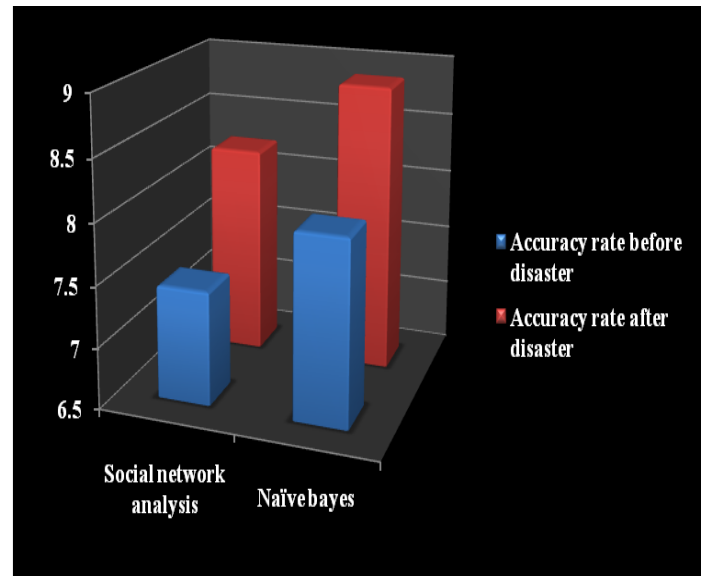

Fig 4.Final report generation

Fig 5.Accuracy graph

Table III
Accuracy Table

Algorithm	Accuracy rate before disaster	Accuracy rate after disaster
Social network analysis	7.43	8.23
Naïve bayes	8.01	8.87

IV. CONCLUSION AND FUTURE ENHANCEMENT

In this system a web application is designed for reporting and sending alert messages when a disaster occurs by using the datasets collected from a social media. The collected data are then analyzed using a data analyses algorithm and then accordingly a graph is generated and an alert message is sent. Here a model of a social media is created for collecting the datasets. The datasets are the tweets that are posted by the user and the tweets are then extracted from the database by using the naïve bayes algorithm and then the collected dataset is then cross verified with a central satellite database and then if the data match with one another an alert message is sent to the users those who are residing in the particular disaster location. The final module of this paper gives the representation of the number of posts posted by the users coming from a specific location and the graph is drawn accordingly with the number of tweets and the location from where the tweets are posted. This model can be further improved. This could also get improve by authenticating the data extraction on a wild scale. This model is however capable of send alert message this could also be improved with voice call functionality. This model which only comprises of extracting the text which processed to get the valid information which however represented as graphical charts, but it could also get improved by extracting photos and videos with the principle of image processing. The system get

superiorly improved when it advances with the feature of artificial intelligence, which gives imaginary character representation like siri and Google assistance, which improves fluidic quality of the application. One of important future enhancements is bringing out this application as a mobile app like other social medias.

REFERENCES

1. Alex Altman. (2015). Why Terrorists Love Twitter. Time. Retrieved from http://time.com/3319278/isis-isil-twitter/
2. Berger, J., & Morgan, J. (2015). The ISIS Twitter Census. Retrieved from http://www.brookings.edu/~/media/research/files/papers/2015/03/isis-twitter-censusberger-morgan/isis_twitter_census_berger_morgan.pdf
3. Chatfield, A. T., Reddick, C. G., & Brajawidagda, U. (2015). Tweeting Propaganda, Radicalization and Recruitment: Islamic State Supporters Multi-sided Twitter Networks. Proceedings of the 16th Annual International Conference on Digital Government Research (dg.o 2015), 239–249. http://doi.org/10.1145/2757401.2757408
4. Johoo Kim & Makarand hastak. Social network analysis: Characteristics of online social networks(2018) International Journal of Information Management. https://www.researchgate.net/publication/322175764.
5. Wu Hea, Shenghua Zha , Ling Li. Social media competitive analysis and text mining: A case study in the pizza industry 2013. https://sci-hub.tw/https://doi.org/10.1016/j.ijinfomgt.2013.01.001.
6. Social media application using jsp,mysql,java. http://studygyaan.com/projects/java/social-media-application-using-jsp-java-mysql.
7. SAP PO – MySQL JDBC connectivity. https://blogs.sap.com/2018/04/20/sap-po-mssql-jdbc-connectivity-issue/.
8. Qiong Wang, George M. Garrity, James M. Tiedje, James R. Cole Naïve bayesian classifier for rapid assignment of rrna sequences into the new bacterial taxonomy https://aem.asm.org/content/73/16/5261.short
9. Geeks for Geeks https://www.geeksforgeeks.org/naive-bayes-classifiers/.
10. Bhanu Pratap B V Get alert on mobile using way2sms http://pyritetechnologies.blogspot.com/2013/04/get-mail-alert-on-mobile-using-way2sms.html.
11. Rui Wang, Yasset Perez-Riverol, Henning Hermjakob, Juan Antonio Vizcaíno Open source libraries and frameworks for biological data visualisation: A guide for developers 2015 https://onlinelibrary.wiley.com/doi/full/10.1002/pmic.201400377.

AUTHORS PROFILE

Dr.P.Tamijiselvy works as a professor in Sri Krishna College of Technology in the Department of Computer Science and Engineering with work experience of 21 years. Her areas of interests are image processing, data mining, and artificial intelligence. Her research works include medical imaging.

V.Suriya Prakash, UG student in Sri Krishna College of Technology in the Department of Computer Science and Engineering.

S.Shriram, UG student in Sri Krishna College of Technology in the Department of Computer Science and Engineering.

N.Vimalesh, UG student in Sri Krishna College of Technology in the Department of Computer Science and Engineering

M.Anitha, Full-Time Ph.D Scholar, in Sri Krishna College of Technology in the Department of Computer Science and Engineering. Her areas of interest are image processing and data mining,