



ALERT USING SENTIMENTAL ANALYSIS OF TEXT MESSAGES

Mrs.R.Priya¹, Ms.J.Abinaya², Ms.K.Kaviyaa³, Ms.M.Keerthana⁴

¹Associate Professor, Anjalai Ammal Mahalingam Engineering College, Kovilvenni, Tamilnadu, India

^{2,3,4}B.Tech/IT., Anjalai Ammal Mahalingam Engineering College, Kovilvenni, Tamilnadu, India

Abstract— At the present time, youth is facing different types of emotions. It may be due to the environment and their surroundings. It is being difficult to communicate with family members and others because of the reserved nature and busy schedule. Social media has become a medium for communication via text messages, images, etc. so it is used by people to share their feelings at most for personal. Therefore, a chatting message is a very important source of information to study the state of a person. The aim of this project is to read all messages and understand hidden semantic using data mining techniques to observe the state of a person and necessary action can be taken if there are any emergency cases in order to save the life of the subject. This project clearly explains the extraction and classification of text messages and giving an alert to the concerned person if there is any negative content. It can also extract positive and negative messages.

Keywords — Sentimental Analysis, Support Vector Machine, KNN.

I. INTRODUCTION

Data mining is a process of finding useful information from a large volume of data sets involving methods of machine learning, database systems, and statistics. There are many applications in data mining such as Data Mining Analytics in Sales/Marketing, Banking/Finance, Health Care and Insurance, Transportation, Medical, Education, Research Analysis. The best way to know about the sentiment of a person is by applying data mining techniques to the text messages a person sends. Applying data analysis techniques to the text messages arrives from the people their sentiments must be known and inferred. The best way to know about the sentiment classify as positive, neutral, and negative of a person is by applying data mining techniques to the text messages a person sends. If a person shows any negative or miss behaviour, activities informing the expressive part of any text message convey the real essence of the conversation between the two counterparts. Hence, it is of prime importance to analyse the emoticons used in any text message so that the real sentiment of the text is accessible.

The goal is to acquire information from the text messages of the user and it is used for various purposes such as sentiment analysis. The model also includes the analysis of emoticons in order to completely parse the statements.

II. PRELIMINARIES

They proposed the method for deciding plans and extracting the information from the huge set of data. it uses the concept of data sentiment analysis for mining knowledge from a large volume of data. Sentiment analysis is otherwise called as opinion mining. This is the field of study which analyses people's sentiments, notions, and emotions towards entities. These entities may be a thing, surveys of people, issues, and films. Social sites such as Face book, Twitter in which text messages show people's status or sentiments. People comment on their Face book account concerning any correct subject of their



consideration [1].

With the advent of Web 2.0 various platforms like Face book, Twitter, Linked-In, Instagram allows citizens to share their comments, views, feelings, judgments on a myriad of topics ranging from education to entertainment. These media platforms contain a large volume of data in the form of blogs, tweets, posts, and status updates, etc. Sentiment Analysis aims to determine the polarity of emotions like happiness, sorrow, grief, hatred, anger, and affection and opinions from the text, reviews, posts that are available online on these platforms. Opinion mining discovers the sentiment from the text based on the given source of text.

It has become complicated due to misspellings, the short form of words, characters repetition, use of native language, slang words, and upcoming emoticons. So each word sentiment is identified appropriately. Sentiment analysis has become a massively active research area and it is studied deeply in data mining. This technique is widely applied in every social and business area because opinions play a vital role in human activities and behaviors [2].

In this paper, they proposed the scope of this research fits in sentiment analysis. The objective of sentiment classification is to classify the opinions of users as positive, negative, or neutral from textual information alone. For that, purpose researchers used data mining classification techniques such as naive Bayes classifier and the Neural Networks. The sentimental analysis distinguished it from the traditional thematic research the way it approaches several issues.

In this paper, an analytical and comparative study of different researches conducted about sentiment analysis in social networks using machine learning is presented. This study analyses in more detail the pre-processing steps which are very important in sentiment analysis process success and are the most difficult especially in the case where the comments are written in a not structured language [3].

In this paper, they present a benchmark setup for analyzing the sentiment with respect to users' medical condition considering the information, available in social media in particular. To this end, it has been crawled the medical forum website 'patient.info' with opinions about medical condition self-narrated by the users.

We restricted ourselves to some of the domains such as depression, anxiety, asthma, and allergy. The focus is given on the identification of multiple forms of medical sentiments which can be inferred from users' medical condition, treatment, and medication. Thereafter, a deep Convolutional Neural Network (CNN) based medical sentiment analysis system is developed for the purpose of evaluation. The resources are made available to the community through the LRE map for further research [4].

This chapter gives a detailed picture of the social media landscape and its users explain the roles of social media in tourism. Significantly, social media impacts from both the consumer and the business perspective are discussed and provided suggestions on social media research opportunities and substantial challenges in terms of methods, data quality, and ethics. It concludes with a call for more refined and systematic approaches to social media research to continuously inform theoretical and practical understandings of their nature and effects [5].



III. PROPOSED FRAMEORK

This proposed framework portrays that it can give an alert if there any emergency cases. And then, it can clearly convey the extraction of positive, negative, and neutral messages.

A. Methodology for proposed work

In this proposed work, we can retrieve an average of chats of each person this can be increased or decreased either by manipulating the code or by increasing the internet speed. The average internet speed is required to execute the code without causing any flaws. The Internet speed required is about 100kbps. Further, the data will be stored in a chat list which can be saved in CSV format. For analysis purposes, it is left off as it is. The data stored in the chat list is in the format of CSV2, which can be simple to write into the file to view in Excel.

B. System architecture

In this system architecture figure 3, we can use two devices to process. The customer has to login to their website before scanning the QR code. After completion of the process, data pre-processing can be applied. The data are extracted from the text messages sent by the user. After the data extracted from the user, it can be stored in the database of the training data set. Then the algorithm can be applied to classify the data.

The SVM can be used to classify the data and the KNN(K- Nearest Neighbor can be applied for classifying the data based on the nearest training sets in the future space. A sentimental analysis algorithm can be used to alert the concerned people.

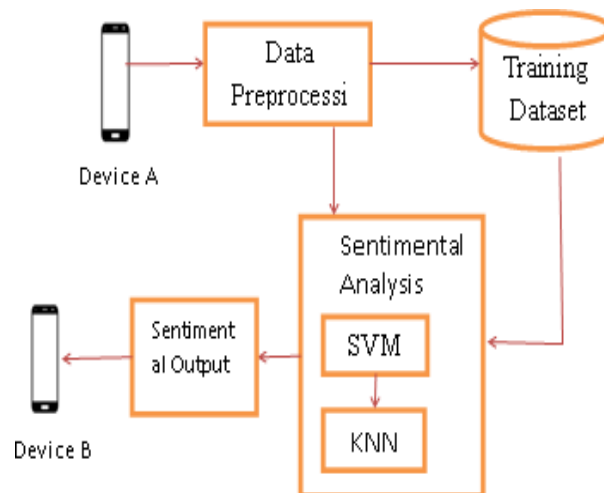


Fig.1 Proposed system architecture



C.Modules and descriptions

Data Preprocessing

The data are extracted from the text messages sent by the user. The analysis requires three different classification results: positive, neutral, and negative. The feature sets for positive, neutral, negative are briefly explained as follows:

- 1) Vocabulary list: The system consists of a vocabulary that includes commonly used words such as sleepy, lazy, etc. Process text messages are checked against this vocabulary and found words are converted into features for emotion recognition.
- 2) Non-meaning full words: Chats also includes non- meaning full words. The system also consists of non- meaning full words like gud , mrng ,enjoy etc.
- 3) Emoticon List: Similar to the vocabulary list, the emoticon list includes commonly used emoticons like happy or sad faces.

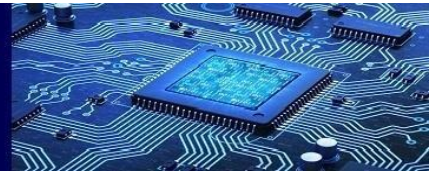
The emoticons are extracted from the text messages and are checked against the emoticon list. Found emoticons are converted into features for emotion recognition.

Training classifies from the training set

The recognition phase next undergoes training. A lot of training examples with their emotional values are fed to the system. When more training sets are used then the system will become better. It trains the classifiers in this phase and it will become ready for actual prediction. As the training data provides the correct sentiment analysis of Emotions of the training examples, using large datasets will ensure more accurate sentiment.

Sentiment prediction using learning algorithm

The resulting stream of words after the text pre- processing step is processed by SVM Algorithm in order to classify the messages as “positive, neutral and negative” sentiment. The process is applied to every message in the data set in order to classify the chat as one among the sentiment. Thus we will get a sentiment associated with the messages associated with the user. SVM are supervised learning models that are used for classification and regression analysis of data used. An SVM model represents examples as points in space; different classes of examples are divided by a certain gap which must be as wide as possible. New examples when mapped into space are predicted to belong to a class of examples based on which side of the gap they fall. The output obtained from Support Vector Machines Algorithm is clusters of two sentiments with class labels “normal” and “critical”. Based on the output KNN algorithm is applied in order to deduce the overall sentiments of the subject. The input for the KNN algorithm is the sentiments associated with all the chats that the subject is involved in. The final step is the sentiment of the person is predicted based on the feature set collected. Data is divided into training and testing sets, and the KNN algorithm is used to predict the sentiment.KNN algorithm is a method for classifying data based on the nearest training sets in the feature space. The class label is assigned the same class as the nearest K instance in the training set. KNN is a type of lazy learner strategy. KNN algorithm is



considered a simple classification technique based on machine learning concepts.

Result Aggregation and Alert user

After the classifiers trained, the actual input may be given to the system. In a particular time span, users can be active on the number of chats. In the end, the individual results are aggregated into a final, and overall sentiment analysis is measured for the user. The system will predict Emotions for every single chat. Finally, all these results are aggregated into a single value of sentiment. These aggregated values of Emotion conclude the sentiment analysis of the user and the alert for a negative message is send to the associated people.

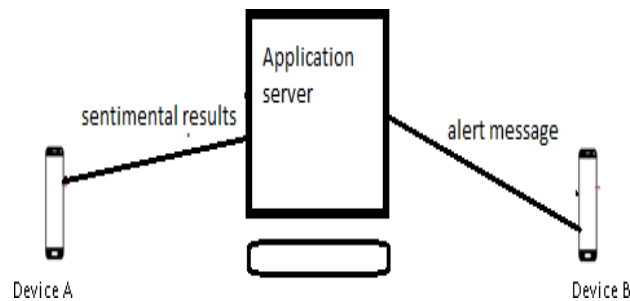


Fig.2 System Architecture for alert user

IV. CHALLENGES & CONCLUSION

A. Challenges

- 1) One of the challenges is to extract the data from the user.
- 2) The internet connection is required.

B. Conclusion

The proposed model takes input from the data set created by accumulating all the text messages sent by the subject. All the messages from WhatsApp.

Keywords are obtained from the data sets after pre- processing the messages. After pre-processing, probabilistic language models are used. Associating weights to the data set increases the overall efficiency of classifying algorithms. The next step is to use the classifying algorithms to classify the conversations as “positive” "neutral “or “negative”. Both a supervised algorithm and an unsupervised algorithm are used to increase the efficiency drastically and prove highly efficient for such computations.



Thus, it is proposed to give a highly efficient method of finding the sentiment of the person by analyzing the text messages and also processing emoticons. Emoticons are very common tokens in any text message.

REFERENCES

- [1] FARHAN LAEEQ ,MD.TABREZ NAFIZ AND MIZRA RAHIL BEG “Sentimental Classification of Social Media using Data Mining,” June 2017.
- [2] PALAKAD BAID,APPORVA GUPTA,NEELAM CHAPLOT “Sentimental Analysis of Movie Reviews using Machine Learning Techniques” December 2017.
- [3] SANDEEP NIGAM, AJIT KUMAR DAS, RAKESH CHANDRA BALABANTARAY “Machine Learning Based Approach To Sentiment Analysis “October 2018.
- [4] SHWETA YADAV, ASIF EKBAL, SRIPARNA SAHA, PUSHPAK BHATTACHARYYA “Medical Sentiment Analysis using Social Media: Towards building a Patient Assisted System” 2018.
- [5] ULRIKE GRETZEL “Tourism and Social Media” January 2018.
- [6] RUPINDER KAUR,DR.HARMADEEP SINGH, DR.GAURAV GUPTA “Sentimental Analysis on Facebook comments using Data Mining Technique”, August 2019.