**SENTIMENT ANALYSIS OF A WHATSAPP CHAT USING R**

**Abstract**

Data, the one, that made a trend and evolution from good old techniques to modern trendy techniques. Based on the source the data is extracted, its form differs that is in recent decades data are scalable, skeptical and diverse compared to old structured data that are derived from specific known sources.The next criterion for the technological evolution is, the storage space that is needed for data. Structured data storage lies in the database whereas for the unstructured and semi-structured ones the storage trend tips to cloud. The vital thing in addition to storage is data handling and analysis, both process when involved in trendy data results in a complexity in terms of containment, processing and visualization of data. To overcome these complexities on trendy unstructured data there emerged a concept called Big Data Analytics. It uses numerous tools and techniques that resolved the problem from data storage to visualization. One such technique is Analytical sandbox, the most central container that stores and handles data in a very appropriate manner. In addition, Big data tools provide a variety of analytical and visualization techniques that produces efficient graphical view of modern data. This paper centered on sentimental analysis of WhatsApp group data in which text mining was incorporated to the chat file and the resulted chat text file is further processed with analytical tools that analyze the chat contents and produces an graphical visualization of the sentiments shared in the group. Its outcome view lies in high echelon of positive opinions compared to other emotions such as anger, fear, disgust, anticipation, joy, sadness, surprise, trust and negative

**Literature review**

There are different methods used for sentiment analysis, including training a known dataset, creating your own classifiers with rules, and using predefined lexical dictionaries (lexicons). In this project, we will use the lexicon-based approach.There are softwares like MS- excel being used for the data analysis but for more powerful and other features .

RStudio is the most favoured IDE for R is been used to perform exploratory data analysis and visualization for the collected data largely because of its open source nature.

We shall use the whatsapp group chat data and personal data to load, clean and perform necessary operations using R functions fir tha analysis.

**SENTIMENTAL ANALYSIS OF WHATSAPP DATA**

Sentimental analysis is the emerging concept which encloses the opinion and attitude of a person. It is interfaced with some techniques and tools through which mining of relevant data is done and emotional or opinion results are produced. It centers on the opinion of a person whether it is negative or positive on certain topic of discussions in a group or team. This paper considers a discussion forum that is WhatsApp group chat for certain time period that analyze the sentiments of the group through chats that are mined from WhatsApp chat group.

**IDENTIFYING AND FORMULATING A PROBLEM**

More than 34 billion texts are exchanged over the WhatsApp every day and just imagine if we could analyze and get valuable insights from this data and leverage it to not only take better real-time decisions but also add value to the stakeholders at much lower cost and time and hence align our operational efficiency with organizational strategy. In this article, we’ll leverage the power of sentiment analysis to investigate the WhatsApp chat using R, visualize and interpret the results at the same time.

WhatsApp is most popular chat app with monthly active users of more than 700 million. The popularity of this app has made it a necessary app among smartphone users and even businesses and organizations use WhatsApp for daily communication in groups and across departments. Corporations get a huge amount of textual data from WhatsApp and they can leverage WhatsApp chat sentiment analysis to gain better insights about their employees and try to avoid unforeseen conflicts due to various redundancies and inefficiency of business processes.

**ANALYSIS STEPS ARE AS FOLLOWS**

* R-tool has an package termed **syuzhet** which is imported at the start into

the R-environment.

* WhatsApp chat text file that is saved¬ in systems local drive is imported

into the environment

* Analysis on the emotions are made¬ using in-built function that

produces sentimental values for the chat file

* Dataframe and colsums are the one¬ that adds additional framework for

the analysis produced.

* Data visualization techniques such as barplot, boxplot, histogram,etc.,that

is suitable for visualizations of the sentiments can be used and this visualization is the final step for exposing analytical result.



**SENTIMENT ANALYSIS PROCESS FLOW MODEL**

**Problem Statement**

Every data science project needs to have a set of questions to explore. Here are a few to keep in mind as we work through this project: is it possible to write a program to determine the mood expressed in a Whatsapp Chat? Are predefined lexicons sufficient? How much data preparation is necessary? What kind of Visualization to use so that we could figure out the correct mood of the chat to determine what of of decisions to be taken on the basis of the results.

**SENTIMENT ANALYSIS OF WHATSAPP CHAT USING R**

**REPORT**

**Submitted By**

**RA1511008010644 Shubham Singh RA1511008010620 Shilpa Suryawanshi**

**RA1511008010654 Sidhant Iyer**

**In partial fulfillment for award of degree**

**Of**

**BACHELOR OF TECHNOLOGY**

IN

INFORMATION TECHNOLOGY

