1. **INTRODUCTION**

Sentiment analysis or opinion mining is useful for review of movies, products, customer services, opinion about any event etc. This helps us to decide whether specific item or service is good/bad or preferred or not preferred. It is also useful to identify opinions of people about any event or persons and also finds polarity of text whether positive, negative or neutral. Sentiment analysis is a type of text classification which can classify text into different sentiments.

1. **RELATED WORK**

Hybrid classification technique has been used for sentiment classification of movies reviews. Integration of different feature sets and classification algorithms such as Naïve Bayes, Genetic algorithm has been carried out to analyze performance on the basis of accuracy. The output of research works shows that hybrid NB-GA is efficient and effective than base classifier and comparing in NB and GA, GA is more efficient than NB. Polarity of document is also an important aspect in text mining. Future engineering with tree kernel has been discussed by [2]. This technique gives better result than other techniques. In the paper author has define two classification models namely 2-way and 3-way classification.

**III.SENTIMENT ANALYSIS**

*A. Dataset Creation*

Dataset was created from retrieval of Uri Attack tweets. To retrieve tweets a tweeter application was generated to get ConsumerKey, Consumer\_Secret, Access\_Token and Access\_Token\_Secret. These keys are used to connect R Studio and tweeter application. Once the connection is done, providing search term as “Uri Attack” a dataset of 5000 tweets was created. This data set was pre-processed to eliminate duplicate tweets so

final dataset contained 1788 tweets

*B. Data Pre-processing*

The final dataset consisted of raw tweets, which needed pre-processing to get good results. Tweets were processed to remove stop words, frequent usage words such as conjunctions, numbers, prepositions, names, base verbs, etc. These type of words do not play any important role in sentiment analysis. Following are preprocessing steps.

 Filtering-In this step tweets are cleaned by removing inks, some special words, emotions symbols, user

names etc.

 Tokenization-In this step tweets are separated into different tokens.

 Stopwords Removal-Stop words are nothing but specific common words which have no analytic value

are removed from the tweets.

 stemDocument-This step is used to remove common word endings such as “ing”, “es”, “s” etc.

 Remove White Spaces-Each text contains lots of white spaces. In this step white spaces are removes.

 Convert to lower-After removing all unnecessary .



**IV.OBSERVATIONS**

Present study classifies tweets into six different emotions namely anger, disgust, fear, joy, sadness and surprise. Also we classified these tweets into three polarities namely positive, negative and neutral. Table 1 shows count of sentiment for different sentiments and table 2 shows polarity count for each type of polarity.Figure 1 and figure 2 shows sentiment analysis about Uri attack according to emotions and polarity. From table 1, it is observed that most of the people have fear about this event. Fear count was highest among all the emotions which are equal to 55.59%. Table also shows that anger and disgust emotions are little bit

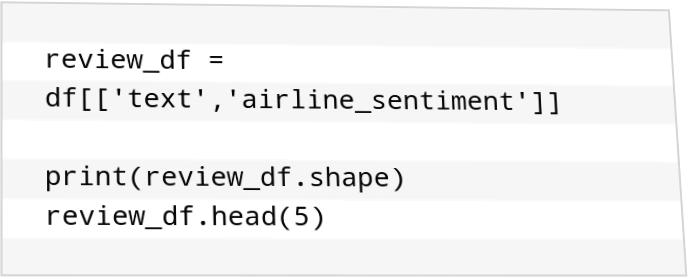
similar and are 18.46% and 19.63% respectively. Some tweets are related to sadness and are equal to 0.62%.

Table also shows that some people were surprised about Uri attack. But shameful fact is that 4.75% people were in joy about Uri attack. This group of people may be terrorist or terrorist supporter. Overall we say that near about 94.3% people have fear, anger, sadness and disgust emotions, and as 5.7% people were surprised and joyed about Uri attack.Polarity of tweet’s aim is to identify overall conceptual polarity of writer. Table 2 shows polarity of the

tweets. From table 2, we observe that 67.17% tweets express negativety, 14.93% express neutral stance and 17.90% were positive.According to human psychology, any attack or terrorist activity generates panic. Present work accurately.















**V. CONCLUSION**

Consumer marketing data is used for collecting sentiments about product and collected data is used for

future prediction. Consumer review data is huge amount of data so author uses hadoop environment for

sentiment analysis. Experimental work created hadoop clusters for analysis of data. Tweets were categorized as

positive, negative and neutral [6].

Hadoop’s FLUME and HIVE tools are also used for analysis of twitter data. FLUME tool extracts data

and stores into HDFS form. HIVE tool is used to extract and analyze data from HDFS type storage. HIVE tool

is helps in analysis of different topics by changing keywords. Author identifies sentiments and polarity of tweets

from election voting data [7].