Indian J.Sci.Res. 17 (1): 329-333, 2017 ISSN: 0976-2876 (Print)
ISSN: 2250-0138 (Online)

EMOTIONAL ANALYSIS OF TWEETS RELATED TO HDFC BANK

PIYUSH BHARDWAJ[a1], RAMESH KUMAR[b] AND PAYAL PAHWA[c]

[ac]Department of Computer Science and Engineering, BPIT, GGSIPU, Delhi, India
[b]Associate Professor, Department of CSE, OPJS University, Churu, Rajasthan, India

ABSTRACT

Emotional Analysis has become a wide research area due to its wide application domain. In this paper, Emotional analysis of tweets related to HDFC bank is done. Banking is one of the biggest service industries in the world. Every person needs to deposit and collect cash from banks. Apart from this, a bank provides many other services like Net Banking, Fixed deposit, NEFT, IMPS etc.. Our approach is applied on one of the biggest private banks in India, HDFC bank. The analysis is done in statistical package R. This system analyses various emotions: JOY, SURPRISE, SADNESS, ANGER, DISGUST and FEAR for emotional analysis. A comparison of emotions is done for various cities of India for HDFC bank. With an analyses of these feedback, a judgement is generated for HDFC as a trustworthy bank.

Keywords: Emotional Analysis, Quality Parameters, Naive Bayes, Service

Emotional examination is the ponder of people's opinions, emotions and criticisms. An arrangement that could incorporate whatever event, subject sentence alternately unique. Emotional analysis will be a territory for investment for a lot of people and scientists with the expanding volume of information.

In (Montejo-Ráez *et al.*, 2014) a technique for class of Emotional Polarity in Twitter posts is done. The author extracts a vector of weighted nodes from the graph of Word Net. The approach proposes a domain-unbiased non-supervised solution for Emotional polarity classification. In the paper (Clavel and Callejas, 2015), the writer identifies and discusses the possibilities that are developing for pass-disciplinary work, so as to enhance man or woman advances. The author gives s comparative analyses of the Emotional detection methods and Emotional-related phenomena utilized in both groups and gives a top level view of the aims of socio-affective human-agent strategies. In (Tsytsarau and Palpanas, 2012), A review of development of Emotional Analysis and Opinion Mining for the duration of past few years is carried out. The paper discusses the evolution of Contradiction Analysis, a brand new research directive. A review of recent advances inside the areas is given and a new approach is proposed for Contradiction Analysis.

The author presents a text analysis model (Malandrakis *et al.*, 2013), which estimates and combines scores of multi-word terms. An application is likewise made to solve the hassle of sentence polarity/semantic orientation detection. The authors started out from a hierarchical compositional technique to generate sentence ratings, and then improved their version by using including multi-word terms, which captures non-compositional semantics.

The author (Subasic and Huettner, 2002) proposes a blend of natural language processing and fuzzy logic techniques to research the affect content material in loose textual content. They show diverse factors of affect analysis with the aid of the usage of news content and numerous movie evaluations. Their effects give an awesome correlation between human judgments of have an effect on content and have an effect on sets.

In (Hua *et al.*, 2016), The author makes a prototype gadget for quick text information. The creator makes use of semantic understanding provided through an understanding base and harvested from a web corpus. Their understanding-in depth procedures unsettle the conventional strategies for acting tasks together with component-of-speech tagging, idea labelling and text segmentation.

The survey (Medhat *et al.*, 2014) provide complete review of sentiment evaluation techniques and few associated fields. The paper consists of a categorization of many latest articles and direction of research within the sentiment evaluation.

In (Martín-Valdivia *et al.*, 2013), the author proposes a model which uses meta-classifiers to combine supervised and unsupervised mastering to make a polarity type machine. The statistics set used is a Spanish corpus of movie evaluations with its English translation. Their device yields higher overall performance in comparison to person systems and confirmed that this technique can be reviewed for polarity category.

Survey (Egozi *et al.*, 2011) covers procedures and strategies for immediately enabling opinion-orientated information-searching for structures. The creator gives an outline of methods that discover new

[1]**Corresponding author**

issues via sentiment-aware packages. A few papers on issues regarding privacy, manipulation, economic impact and evaluative text also are protected.

In (Montoyo *et al.*, 2012), the writer affords a top level view of research in Emotional analysis and subjectivity and, their utility domains in Natural Language Processing. The author provides few achievements acquired so far and the troubles that stay to be handled.

In paper (Favaretto *et al.*, 2006), the author has formulated a mathematical formula of the temporal–Traveling Salesman Problem with Time Windows. A meta-heuristic based on Ant Colony System is likewise proposed and carried out. A Computational enjoy on a benchmark hassle is supplied and effects are received and analyzed on the premise of a case observe.

The paper is organised as follows: Section 2 explains the methodology used analysis. Section 3 provide the results by giving various parameters showing the comparative results of various cities. Section 4 concludes the paper followed by future scope in Section 5.

METHODOLOGY

We have defined our approach for calculating the emotions of a text. For emotional analysis, we define six classifier sets as follows:

The probability of an angry tweet (E1), a disgusting tweet (E2), a fearful tweet (E3), a joyful tweet (E4), a sad tweet (E5) and a surprise tweet (E6). All thesesix sets are mutually exclusive and exhaustive sets. All these six sets are defined as:

For calculating the emotion of an angry tweet, we have used the equation:

$$\mathrm{P(Angry \mid Tweet)} = \frac{\mathrm{P(Angry)P(Tweet \mid Angry)}}{\mathrm{P(tweet)}} \qquad (1)$$

Where P(Angry | Tweet) is the probability of an angry tweet. The probability of $\mathrm{P(Tweet \mid Angry)}$ is calculated as:

$$\mathrm{P(Tweet \mid Angry)} = \mathrm{P(w1 \mid Angry)} + \mathrm{P(w2 \mid Angry)} + \cdots + \mathrm{P(w_n \mid Angry)} \qquad (2)$$

Where w1 to $\mathrm{w_n}$ are all the words in the tweet.

For calculating the emotion of a disgusting tweet, we have used the equation:

$$\mathrm{P(Disgust \mid Tweet)} = \frac{\mathrm{P(Disgust)P(Tweet \mid Disgusr)}}{\mathrm{P(tweet)}} \qquad (3)$$

Where P (Disgust | Tweet) is the probability of a disgusting tweet. The probability of $P(Tweet \mid Disgust)$ is calculated as:

$$P(Tweet \mid Disgust) = P(w1 \mid Disgust) + P(w2 \mid Disgust) + \cdots + P(w_n \mid Disgust) \qquad (4)$$

Where w1 to w_n are all the words in the tweet.

For calculating the emotion of a fearful tweet, we have used the equation:

$$P(fear \mid Tweet) = \frac{P(fear)P(Tweet \mid fear)}{P(tweet)} \qquad (5)$$

Where P(fear | Tweet) is the probability of a fearful tweet. The probability of $P(Tweet \mid fear)$ is calculated as:

$$P(Tweet \mid fear) = P(w1 \mid fear) + P(w2 \mid fear) + \cdots + P(w_n \mid fear) \qquad (6)$$

Where w1 to w_n are all the words in the tweet.

For calculating the emotion of a joyful tweet, we have used the equation:

$$P(Joy \mid Tweet) = \frac{P(Joy)P(Tweet \mid Joy)}{P(tweet)} \qquad (7)$$

Where P(Joy | Tweet) is the probability of a joyful tweet. The probability of $P(Tweet \mid \mathrm{Joy})$ is calculated as:

$$\mathrm{P(Tweet \mid Joy)} = \mathrm{P(w1 \mid Joy)} + \mathrm{P(w2 \mid Joy)} + \cdots + \mathrm{P(w_n \mid Joy)} \qquad (8)$$

Where w1 to $\mathrm{w_n}$ are all the words in the tweet.

For calculating the emotion of a sad tweet, we have used the equation:

$$\mathrm{P(Sad \mid Tweet)} = \frac{\mathrm{P(Sad)P(Tweet \mid Sad)}}{\mathrm{P(tweet)}} \qquad (9)$$

Where P(Sad | Tweet) is the probability of a sad tweet. The probability of $\mathrm{P(Tweet \mid Sad)}$ is calculated as:

$$\mathrm{P(Tweet \mid Sad)} = \mathrm{P(w1 \mid Sad)} + \mathrm{P(w2 \mid Sad)} + \cdots + \mathrm{P(w_n \mid Sad)} \qquad (10)$$

Where w1 to $\mathrm{w_n}$ are all the words in the tweet.

For calculating the emotion of a surprise tweet, we have used the equation:

$$\mathrm{P(Surprise \mid Tweet)} = \frac{\mathrm{P(Surprise)P(Tweet \mid Surprise)}}{\mathrm{P(tweet)}} \qquad (11)$$

Where P(Surprise | Tweet) is the probability of a surprise tweet. The probability of $\mathrm{P(Tweet \mid Surprise)}$ is calculated as:

$$P(\text{Tweet} \mid \text{Surprise}) = P(w1 \mid \text{Surprise}) + P(w2 \mid \text{Surprise}) + \cdots + P(w_n \mid \text{Surprise}) \quad (12)$$

Where w1 to w_n are all the words in the tweet.

RESULTS AND ANALYSIS

The model is executed in statistical Package R. It is one of the most powerful tool for data analysis. R was created by Ross Ihaka and Robert Gentleman at University of Auckland, New Zealand, and is currently developed by the R Development Core Team.

The emotional analysis is done on tweets related to HDFC bank from the cities Bangalore, Chennai, Delhi, Gurgaon, J&K, Mumbai and Noida. A detailed analysis can be seen in fig 1 and fig 2. Average emotions are calculated and shown in table 1. A detailed analysis of emotions is shown in fig 3 and 4.

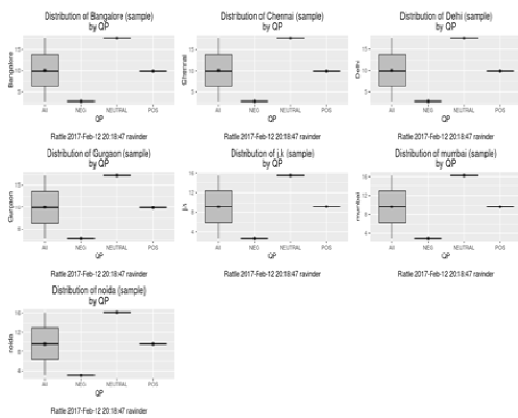

Figure 1: Showing the comparison of cities Bangalore, Chennai, Delhi, Gurgaon, J&K, Mumbai and Noida against quality parameters

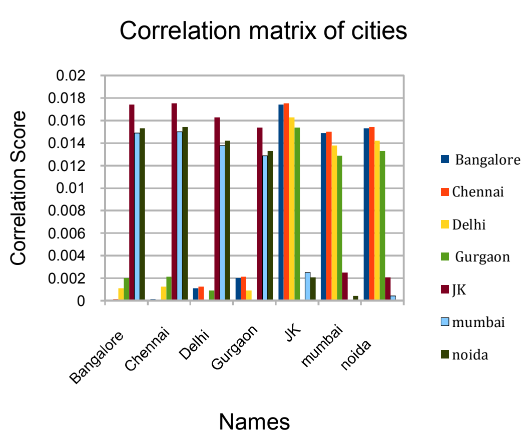

Figure 2: Showing correlation among cities against Emotional scores

Fig 2 presents the correlation among cities and Emotions. Bangalore and chennai shows high correlation against each other. Delhi, gurgaon and noida also shows high crrelation among each other. Mumbai and Jammu & Kashmir shows moderate correlation score of 0.003 among each other

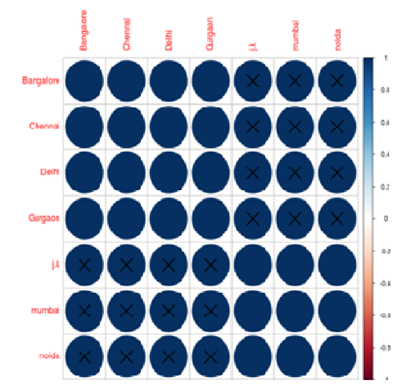

Figure 3: Showing the Correlation map between cities Bangalore, Chennai, Delhi, Gurgaon, J&K, Mumbai and Noida against quality parameters as mentioned in table (1)

Fig 3 gives the analysis of various cities using correlation. The values for correlations are represented by the correlation coefficients (C). The range of possible values for (C) varies from -1.0 to +1.0. The correlation coefficients (C) having values less than zero represents a negative relationship while greater than zero represent a positive relationship.

Table 1: Showing Emotional analysis results of each city

S.No	City	ANGER	DISGUST	FEAR	JOY	SADNESS	SURPRISE	POS	NEG	NEUTRAL
1	Bangalore	3.52	3.41	2.09	3.1	2.06	4	9.93	2.8	3.57
2	Chennai	3.45	3.41	2.09	3.1	2.06	3.97	9.84	2.9	3.45
3	Delhi	3.57	3.4	2.12	3.1	2.06	3.88	10.2	3	3.4
4	Gurugram	3.57	3.4	2.12	3.1	2.06	3.89	10.2	3	3.39
5	j&k	3.84	3.47	2.15	2.7	2.08	3.99	10.5	3	3.49
6	Mumbai	3.88	3.4	2.12	2.7	2.1	4.06	10.6	2.9	3.68
7	Noida	3.57	3.4	2.12	3.1	2.06	3.88	10.2	3	3.4

Table 1 shows the Emotions of people of all the cities used for analysis. The table shows that Emotion surprise has a maximum value among all Emotion followed by anger, disgust, joy, sadness and fear.

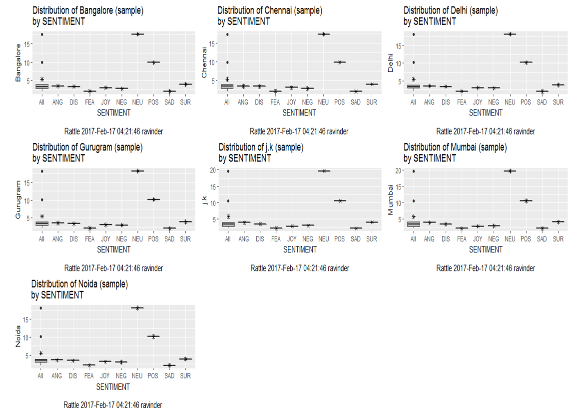

Figure 4: Showing the distribution for Emotions against cities Bangalore, Chennai, Delhi, Gurgaon, J&K, Mumbai and Noida

Fig 4 presents the distribution of Emotions against cities. This gives a pictorial representation of all the emotions and opinions yielded from the system.

CONCLUSION

Emotional Analysis is used in a variety of areas ranging for decision making. The proposed system has an accuracy of more than 87 percent thus making it superior from many other techniques. The system is used in Banking industry. The system analyses the emotions for one of the biggest bank in India: HDFC. The system has compared seven different cities in India against the parameters JOY, SURPRISE, SADNESS, ANGER, DISGUST and FEAR. The maximum value of emotions is scored by emotion SURPRISE with a highest value of 4.06. This shows that most of the cities have a positive opinion and a surprise Emotional for HDFC. Almost all the cities are satisfied with the company. Thus we conclude that HDFC is a trustworthy company and people can rely on the service for their financial needs

FUTURE SCOPE

The proposed system will be used to review various banks in India. A comparative analysis will be done to retrieve the most trustworthy service provider in the India.

REFERENCES

Cambria E., Schuller B., Xia Y. and Havasi C., 2013. New avenues in opinion mining and Emotional analysis. IEEE IntellSyst, **28**: 15–21.

Clavel C. and Callejas Z., 2015. Emotional analysis: From opinion mining to human-agent interaction. IEEE Transactions on Affective Computing, **7**(1): 74-93, from http://ieeexplore.ieee.org/document/7122903/

Egozi O., Markovitch S. and Gabrilovich E., 2011. Concept-based information retrieval using explicit semantic analysis. ACM Transactions on Information Systems (TOIS), **29**(2): 8. doi:10.1145/1961209.1961211

Favaretto D., Moretti E. and Pellegrini P., 2006. An ant colony system approach for variants of the traveling salesman problem with time windows. Journal of Information & Optimization Sciences, **27**(1): 35-54.

Feldman R., 2013. Techniques and applications for Emotional analysis. Commun ACM, **56**:82–9.

Hua W., Wang Z. and Wang H., 2016. Understand short texts by harvesting and analyzing semantic knowledge. IEEE Transactions on Knowledge and Data Engineering, **29**(3): 499-512, from http://ieeexplore.ieee. org/document/7476863/

Liu B., 2012. Emotional analysis and opinion mining. Synth Lect Human Lang Technol.

Malandrakis N., Potamianos A., Iosif E. and Narayanan S., 2013. Distributional semantic models for Affective text analysis. IEEE Transactions on Audio, Speech, and Language Processing, **21**(11): 2379-2392, from http://ieeexplore. ieee.org/document/6578101/

Martín-Valdivia M.-T., Martínez-Cámara E., Perea-Ortega J.-M. and Ureña-López A.L., 2013. Emotional polarity detection in Spanish reviews combining supervised and unsupervised approaches. Expert Systems with Applications, **40**(10): 3934–3942. doi:10.1016/j.eswa. 2012.12.084

Medhat W., Hassan A. and Korashy H., 2014. Emotional analysis algorithms and applications: A survey. Ain Shams Engineering Journal, **5**(4): 1093–1113, from http://www.sciencedirect.com/ science/article/pii/S2090447914000550

Montejo-Ráez A., Martínez-Cámara E., Martín-Valdivia T.M. and Ureña-López A.L., 2014. Ranked WordNet graph for Emotional polarity classification in Twitter. Computer Speech & Language, **28**(1): 93–107. doi:10.1016/j.csl. 2013.04.001

Montoyo A., Martı´nez-Barco P. and Balahur A., 2012. Subjectivity and Emotional analysis: an overview of the current state of the area and envisaged developments. Decis Support Syst., **53**:675–9.

Pang B. and Lee L., 2008. Opinion mining and Emotional analysis. Found Trends Inform Retriev., **2**:1–135.

Subasic P. and Huettner A., 2002. Affect analysis of text using fuzzy semantic typing. IEEE Transactions on Fuzzy Systems, **9**(4): 483-496, from http://ieeexplore.ieee.org/document/940962/

Tsytsarau M. and Palpanas T., 2012. Survey on mining sub- jective data on the web. Data Min. Knowl. Discov., **24**:478–514.

Wilson T., Wiebe J. and Hoffman P., 2005. Recognizing contextual polarity in phrase-level Emotional analysis. In: Proceedings of HLT/ EMNLP.