International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-3, January 2020

Assorted Sentiment Model for Publically Available Page of Facebook

Saurabh Dhyani, G. S. Thakur

***Abstract** ecommerce industries expose public page in the social network site (Facebook, twitter etc) for the intention of improving of business strategy. They extract public mood about the social network page in the forms of total likes, the total share of the page and sentiment of all comments to the social network page similar way celebrities expose public page in the social network sites for the intention of improving its fame. We have developed an assorted model for publicly available page of Facebook. This assorted model is the combination of data extractor model, language convertor and cleaned model, and sentiment analyzer model. Our data extractor model extract comments on all the posts of publicly expose Facebook page in the less span of time. Language convertor and cleaned model would work for conversion of text written in different Indian language to the English language and after that English written text would be cleaned through cleaned model. Language convertor is made after implementing CILTEL model. CILTEL model converts comments written in the Indian languages in the English language. Cleaning model will clean all the comments of all the posts on the Facebook page. Finally, sentiment extraction model will extract sentiments of all the comments of the Facebook page. We have implemented classification using three machine learning algorithm, namely naïve bayes algorithm, perceptron algorithm and rocchio algorithm for checking the performance of our sentiment analysis model. Our assorted sentiment analysis model is beneficial to users like marketing industry, election parties and celebrities.*

***Keywords:** Sentiment Analysis, Social media mining, Classification, Facebook, Content Mining, Machine learning algorithm*

Abbreviations: SNP: Social Network Page ; PO: Post ; SC, C: Comment ; MG: Merger ; TF: Text File ; CTF: Cleaned Text File ; BOW: Bag of words ; CILTEL : Converter of Indian languages to English language.

I. INTRODUCTION

From Starting of Centaury, Web two has been expressed as an intellectual way and also expressed to improve the technical base for providing growth to the Content generated by user. Online marketing Strategy and Blogging platforms are the first example of technical foundation [1].

Revised Manuscript Received on January 30, 2020.

* Correspondence Author

Saurabh Dhyani *, PhD Degree, Department of Computer Applications, Maulana Azad National Institute of technology. Bhopal (Madhya Pradesh) India.

Ghanshyam Singh Thakur, Assistant Professor, Department of Computer Applications, Maulana Azad National Institute of technology. Bhopal (Madhya Pradesh) India.

User generated content is growing rapidly due to which interest and activity in sentiment analysis and opinion mining is sparking rapidly [2]. The Intention of Sentiment analysis is to extract sentiment from textual data after performing computational process [3].

Sentiment analysis is studied under the name such as opinion mining [2], point of view [4], and subjectivity [5]. The Main contributor to user generated content or textual data is social media. Diffusion of textual data has created new opportunities to fetch public opinion related to any topic. Review data were used previously for finding sentiment analysis. But now the condition has changed to social media. It is generating past to live data. It is challenging task to find sentiment analysis of online textual content of social media. Sentiment analysis is becoming more popular for social media sites which take account of online registered users who are free to communicate their emotions and feelings related to a specific topic. Few studies related to analysis of sentiment have previously been explored. These studies are targeted to social networking sites just like Facebook. Twitter has been used to update a specific topic, mainly for the branding of the products. Twitter update is required to restrict 140 characters in the length [6]. Studies of usage of twitter proved that around 18% tweets include an orientation towards the product 21% of which proved expression about the sentiment of the brand [7]. Opinion analysis has been implemented to opinionated content like news articles and online reviews. However, social networking site content imposes various distinctive challenges to NLP in broad way and to extract sentiment analysis from it in a particular way [2]. Social media platform imposes some of these challenges which are very in nature such as limited length associated to social networking mining. Further, difficulties which are generated by the distinctiveness of communication among users in the social media e.g. need of propinquity, utilization of particular languages and short attention span. The effect is a kind of content that is full of abbreviations and jargon, short and prevalent by reference to the content of social media for finding sentiment of social media contains different approaches [8, 9, 10]. Content centric has been used as a major technique. Authors have been developed detailed feature of linguistic for social networking sites such as previous research has prepared for movie reviews and news articles .Some task has been done to overcome small texts and abbreviation in social networking site by applying an external source to link of contents of text to Wikipedia pages [11] and news articles [12]. When the contents of the text of the social media is associated with multimedia like videos or images, the sentiment analysis of these kinds of media obtained by multi modal analysis [13]. Information shared in the social media is not isolated. it is the reason of these approaches are failing.

The meaning of the content (blog post, Facebook post, tweet) of social media may be understood properly when its contextual kind of data is taken for concern. The contextual kind of data contains observable information just like a prior contents of the text of the social networking site that is related to the identical conversion, previous communications among users and users that communicated with the content of social media (e.g. by loving it, by sharing it). It also includes isolated features of the content of social media. For example, some factors of Demographic like gender and age have been revealed to show a relationship with vocabulary and sentiment [14]. They have been used to get better the result of classification of sentiment [15].

New sentiment analysis approaches are beginning to associate the synthesis of information from social context to the textual content. Context of social networking site has not only been used to find sentiment analysis but also used to find the detection of spam [16]. in this research paper, we have targeted facebook as a social networking sites for implementing assorted sentiment analysis model. The main reason for is that twitter contains 140 lengths of characters for any kind of post but Facebook contains 5000 characters for every update to the post of status [17]. In the status message, sentence formation would be more expectable in social networking sites just like facebook. Sentiment analysis of social networking site is involved during the study of the opinion of the textual data. Furthermore, sentiment analysis of contents of textual data is frequently used by creators associated to film advertisers and other organization that desire to receive response of clients on a particular subject within social network site. In the system developer's prospective; mining of opinion of textual content of social media in sentiment analysis faces some difficulties as well as some challenges. In the case of sentiment of contents of textual data, not every word in the content of sentence contains importance. Some words of sentence are not utilized in the procedure of classifying the sentiment the opinion of the sentence of textual data because some words are classifying as noisy. Besides of words, symbols of emotion like happy face and sad face represent importance in NLP. Hence, in the observation procedure, there are not only words which take place .assorted sentiment analysis works for social media like Facebook with efficient time complexity compared to the other social media. Work related to it has been done by many researcher but they do not focus on the Indian languages .they focus only on English related text to extract sentiment from textual related content from social media site

II. LITERATURE REVIEW

Sentiment analysis of textual data has been used as the computational kinds of study of opinion, emotions and feelings articulated in the textual data [18] .For the propose of the improvement of a sentiment analyzer model for different Indian languages we have tried to keep away from compound and contentious definition of sentiment and emotion. In this directive we have taken the basic characterization of sentiment as an individual negative, positive and neutral polarity. An example of content of sentence presenting the positivity of sentiment would be "I like study" whereas "it is a very bad incident" represent as a negativity of sentiment. A neutral kind of sentiment does not present any kind of emotions. e.g. "I am playing a game". in this research area most of the work pay attention to find classification to the textual data according to their polarization of sentiment which can be neutral, negative and positive [2]. Sentiment analysis is used mainly for classifying of the textual data into neutral polarity, positive polarity and negative polarity. Over the last two decades various works has been done on analysis of sentiment in different kind of application and sentiment analysis can be found using different techniques, further these techniques can be split into three techniques namely dictionary based technique, machine learning technique and hybrid technique [10]. Among these three techniques, hybrid approaches and machine learning technique seem to be dominant [19], and machine learning technique incorporated with lexicon technique for improving their results. Machine learning technique is used for applying a classifier on the combination of parameters that represent input dataset. Difficulties in these methods are to focus for fetching compound parameter from the textual data, discarding only invalid feature and choosing a optimized predictor [20]. First task of sentiment analysis was on review data of movie by classifying it to positive sentiment and negative sentiment via machine learning algorithms [21]. Sentiment analysis was used as the first work to analyze the contents of textual data written in stock exchange market for finding the sentiment of market [22]. Sentiment analysis carried out on the data gathered from Amazon based on the product reviews [35]. The results of experimentation gave optimized solution for both review level and sentence based level classification. Exploration of product opportunity analyzed after handling of sentiment analysis of textual data of online website in business industry [36]. Now days ecommerce industry uses sentiment analysis of textual data with the help of machine learning algorithm and artificial intelligence technique. Technique utilized in modeling of topic of textual data and sentiment analysis for finding the changing needs of customer. Analysis of sentiment of online textual data was carried out Domains like hotel reviews [37] and pizza industry [38] for analyzing the satisfaction level of the customer. At present, large number the works has been done in the area of sentiment analysis for the pay attention to classification of textual data based on the online content of the website. One of the most well-known areas used that of reviews of textual content of online websites of industries [39]. Review website like Epinions is an example of useful source for sentiment analysis. Many areas in which sentiment analysis can be helpful are: misuse of social media, election monitoring Flame detection [23], business intelligence [24], politics or marketing [25], prediction of negative sources or hostile [26], adaption of daily tools dynamically such as e-mail [27], human-computer interaction [28], receiver sentiment based on the broadcasting [29], the context in which words are used accuracy is strongly influenced by it[30]. For example, the sentence "we should play cricket" is negative if the review is about film and it is positive in respect of sports review. In [31], the authors applied lowest amount cuts in the graph for extracting the parts of textual content. They studied and implemented machine learning algorithm for carrying out sentiment analysis on the portion of the textual content.

In [32], authors provided a good study of various kind of technique implemented during sentiment analysis of textual content of online social website. It focuses on the concept of feelings in the written content. Various techniques which can be generally separated into two classes, namely (i) machine learning technique (ii) techniques related to symbolic.

Symbolic technique focuses on the BOW (bag of words approach). depending on the study, the authors identify that machine learning approach achieve good accuracy of classification than that of symbolic techniques on movie review related data for finding sentiment of textual content. Among the machine learning techniques they used three supervised technique (1) Naive Bayes Multinomial (2) Maximum Entropy (3) Support Vector Machine. Significant efforts for fetching twitter data on sentiment analysis organized by [33]. The authors used prediction of sentiment from three online websites to instruct a online model as a noisy label and use around 1001 manually created labeled based tweets from twitter for training set and another 1001 manually created labeled tweets from twitter for testing purpose. However, they do not state that how they gather their testing purpose data. Authors anticipated the feature of grammar of text of tweet from twitter like hash tags, repeat tweet, expletive marks and punctuation in conjunction with the feature like part of content of speech and polarization. In [5], authors use reaction data from universal sustain services study for performing sentiment analysis propose of their learn was to evaluate the task of linguistic feature like part of speech tags of content of social media. They carry out feature selection and extensive feature analysis and exhibit that linguistic feature contribute to the correctness of classifier. Accuracy of classifier is more focused on the feature extraction technique. These techniques include BOW (Bag of Words) feature, POS (Part of Speech) tagging and lexicon based like Bing Liu, AFINN, and general inquirer

III. PROPOSED MODEL

We have constructed assorted sentiment analysis model for social networking sites like Facebook. it has made into three phases, namely data extractor model, language convertor and cleaned text model, and Sentiment analyzer model with machine learning algorithm converts language of content present in the text file TF to English language. For getting cleaned text file (CTF) file, following steps has taken to clean social media content present in the text file TF.

Data Extractor model: Interested social network page contains large number of posts. Any social network registered user in the world is free to put their views regarding post in the form of comment. In this social media text collector model, we can extract posts and comments from any interested publicly available social network page and with the help of MG(MERGER) we store all the comments related to a single post in individual text file. As shown in figure 1 from single publicly available page we can extract n number of post (PO_1,PO_2,PO_3..........PO_n) each post contains variable number of comment, Post PO_1 contains m number of comments(C_1,C_2,C_3...........C_m), Post PO_2 contains n number of comments and so on.. And lastly PO_n contains k number of comments. MG_1 (MERGER1) merges all the comments related to post PO_1 into text file TF_1. MG_2 (MERGER2) merge all comments related to post PO_2 into text file TF_2 and so on. Lastly MG_n (MERGERn) merge all comments related to post PO_n into text file TF_n.

Language convertor and cleaned text algorithm: As shown in the figure 2 there are n text files (TF_1, TF2........TF_n). GR (GENERATOR) generates one text file from n text file. GR (GENERATOR) merges all the contents of n text file into one text file. Social media Content of text file may have written in different languages. COL (CONVERTER OF LANGUAGE)

Steps involved for cleaning the text of text file

1. Encoding Text in standard encoding Format 'UTF-8'
2. All the apostrophise in the text file converted into standard lexicon
3. Remove stopword in the text from text file
4. Remove all the punctuation mark from text in the text file
5. Remove all the human expressions like crying, laughing etc from text in the text file
6. Disjoint attach word in the text of text file
7. Social network text comprises majority of slang word these words converted to standard word to make free text
8. Standardizing words
9. Spelling checking in the text file
10. Grammar Checking in the text file

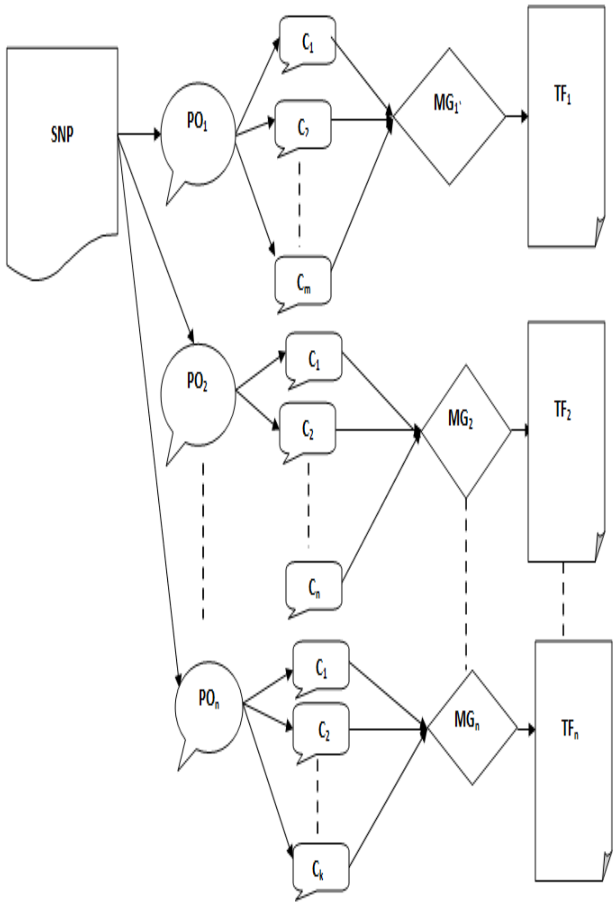

Figure 1: Data Extractor model

CILTEL model for converting Indian languages from one text file to the English language and place English language text to the text file is given in the figure 3. Scanner scans text from the text file and passes the text of the text file to the Indian language detector.

Indian languages detector detects the 10 Indian languages in the written text of the file. Next 10 detected languages are sent to 10 translators translator translate each Indian language to the English language. Translators are synchronized in such a way that at a time only one translator works. example if translator 1 is working for translating Guajarati written text to the English written text that time all other translator will not work.

We move towards next stage when the entire translators have finished work properly. In the next stage placers place English written text to the text file. After performing conversion of languages cleaning algorithm is applied.

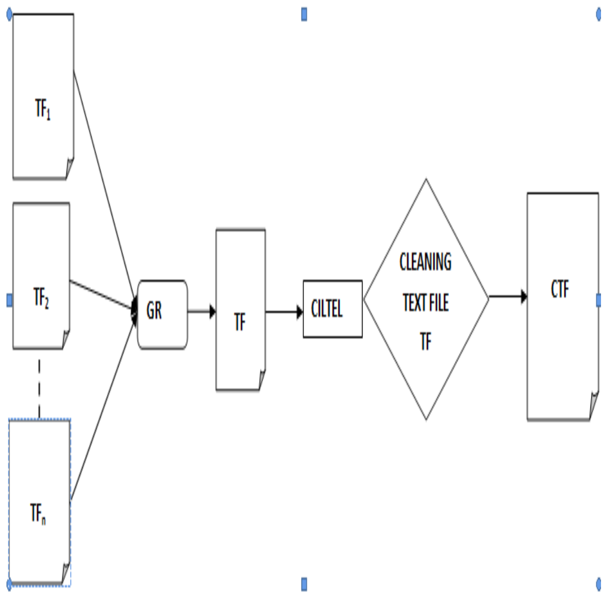

Fig 2: Language convertor and Cleaned model

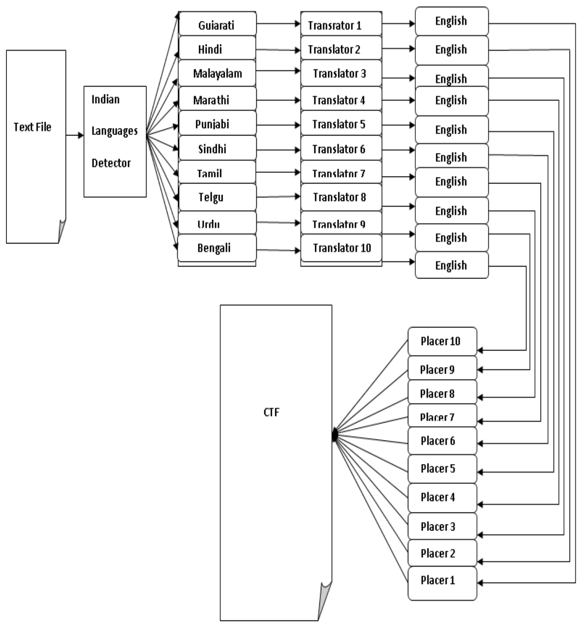

Figure 3: CILTEL Model

Sentiment analyzer model is applied to the cleaned text file obtained from facebook page. Sentiment analyzer model uses machine learning algorithm to find sentiment of the CTF (cleaned text file). Before applying machine learning algorithm we have extracted features like BOW (Bag of Word) feature. For feature extraction, we have used Bing Liu based dictionary or lexicon. Dictionary (Lexicon) of Bing Liu is an opinion based. It is freely distributed and maintained by Liu [3].

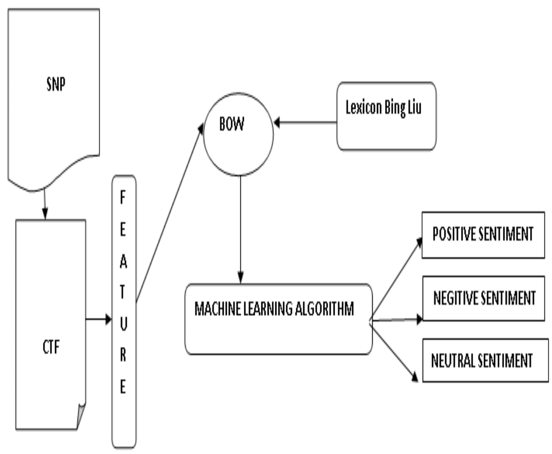

Figure 4: Sentiment Analyzer model

Naïve Bayes Classifiers: it is a collection of classification algorithm based on Bayes' Theorem. Bayes' theorem finds the probability of an event happening given the probability of another event that has already happened. Bayes' Theorem mathematically is expressed as the following equation

$$P(B) = \frac{P(A)P(A)}{P(B)}$$

Where A and B are events P(A) is the priori probability i.e. probability of event before evidence is seen P(A/B) is the posteriori probability of B i.e. probability of event after evidence is happen.

Perceptron Classifier: the perceptron is an algorithm for learning a binary classifier called a threshold function: a function that maps its input a(real valued vector) to an output f(a)(a single binary value)

$$f(a) = \begin{cases} 1, & if\ (w.a + b) > 0 \\ 0, & Otherwise \end{cases}$$

Where w is vector of real valued weights and w.a is the dot product $\sum_{i=1}^{n} w_i a_i$ where n is the number of input to perceptron and b is the bias

Rocchio classifier: text classification using TF and IDF vectors to represent text file. the classifier, which contains nearest point observation to the centriod is identified as the Rocchio Classifier in the technique of machine learning, nearest point observation to the centroid is a classifier model of classification that assign observations to the tag of the group belonging to the training samples whose centroid is the nearest to the observation

IV. RESULT AND DISCUSSION

Python language was used for implementation of data extractor model. From this model we have fetched comments of 4079 posts of Facebook page of prime minister of India (Narendra Modi). Our model has placed comments of 4079 post in the 4079 text files in 20 minutes. While API model of Facebook took 30 minutes for the same task. Extracting number of text files with respect to time from our model is given in the figure 5.

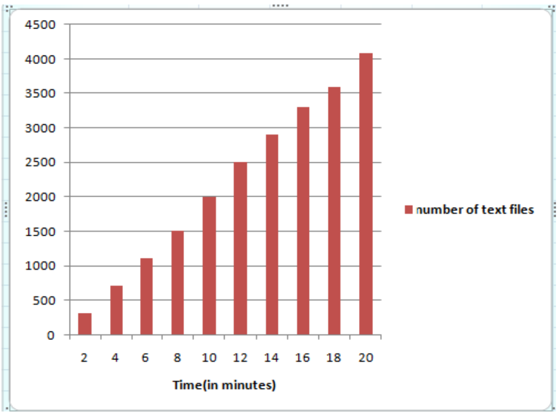

Fig 5: extraction of number of text files with respect to time

We have implemented CILTEL model using Python Script. we have distributed written language in the text file into two stage namely low stage language and high stage language. High stage languages include Hindi, English and low stage languages include other Indian language except Hindi and English. First we have extracted number of comments written in the low stage languages of all the posts of Facebook page through CILTEL model. Figure 6 shows the number of comments written in the low level Indian languages in the all post of the Indian Prime Minister (Narendra Modi) on public Facebook page.

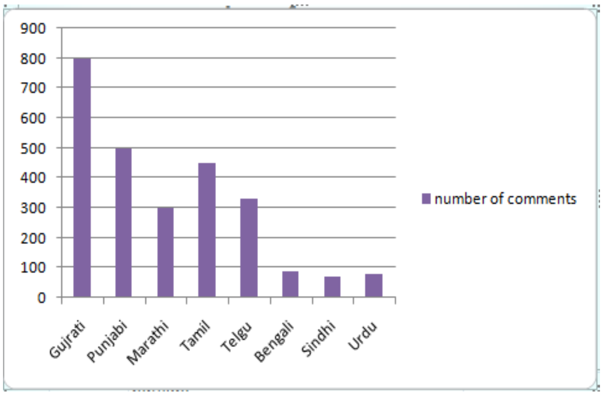

Fig 6: extraction of comments in different low stage Indian languages

Second through CILTEL model we have extracted number of comments written in high stage languages on posts of Facebook public page. Figure 7 shows the number of comments written in the high stage language of all the posts of public page of facebook

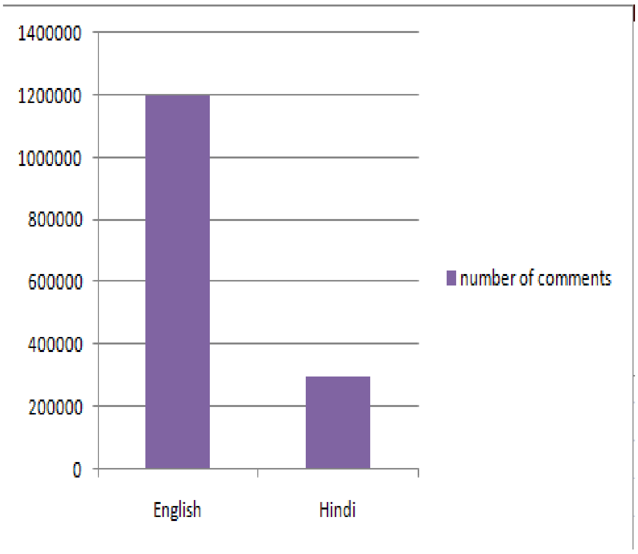

Fig 7: extraction of comments in high stage Indian lang.

Figure 8 shows the time taken by CILTEL model to convert Indian languages to English languages from task file

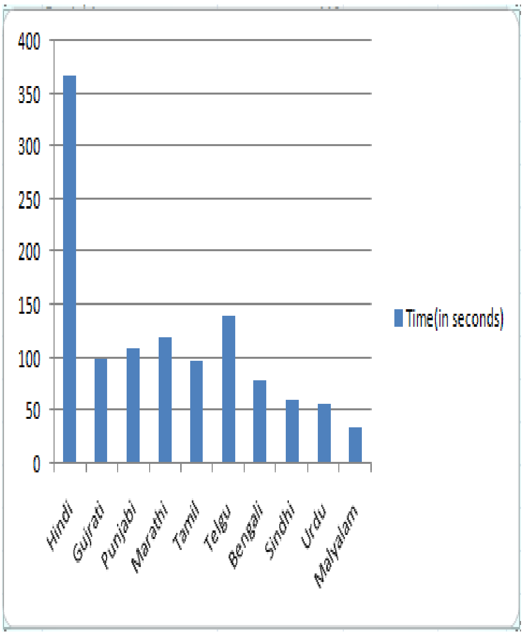

Fig 8: time consuming by CILTEL model in low stage language

After converting all Indian language into the English language in the text file we proceed towards the cleaning for the text present in the text file. During cleaning text file, we have removed emojis, stopwords, punctuations, unusable mark Figure 9 shows distribution of removal feature

Retrieval Number: B7739129219/2020©BEIESP
DOI: 10.35940/ijitee.B7739.019320
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

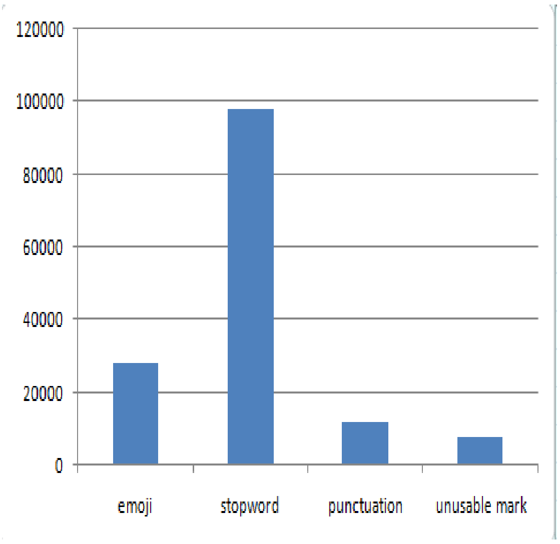

Fig 9: distribution of removal features

Sentiment of the comments of all the posts is extracted through nltk library of python Figure 10 shows the fraction of positivity, negativity and neutrality of sentiment on the comments written in the mixture of different languages on all the posts of page of Indian Prime Minister (Narendra Modi) of facebook. Time taken to extract sentiment before applying pre-processing and CITEL model to Facebook page is 110 seconds

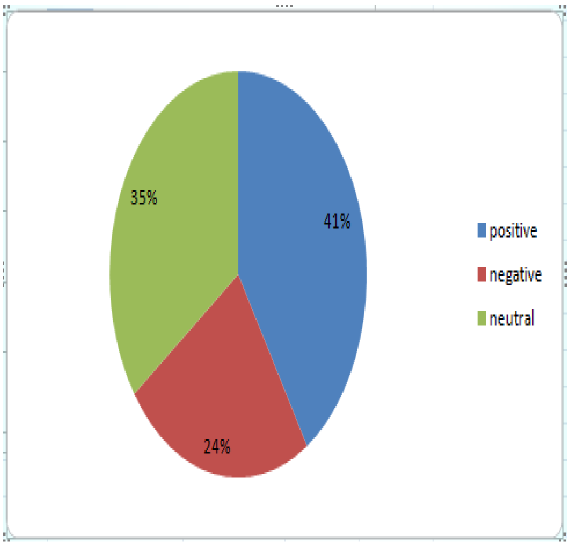

Fig 10: Sentiment of text of mixture of different Indian languages

And CITEL model extracting sentiments takes 70 seconds on one Facebook page. Figure 11 shows sentiment on the all the posts of narendra modi (Indian prime minister) publically Facebook page after converting mixture of different Indian languages to the English languages. Percentage of positive and negative sentiment increased but neutral sentiment is reduced. After applying cleaning model

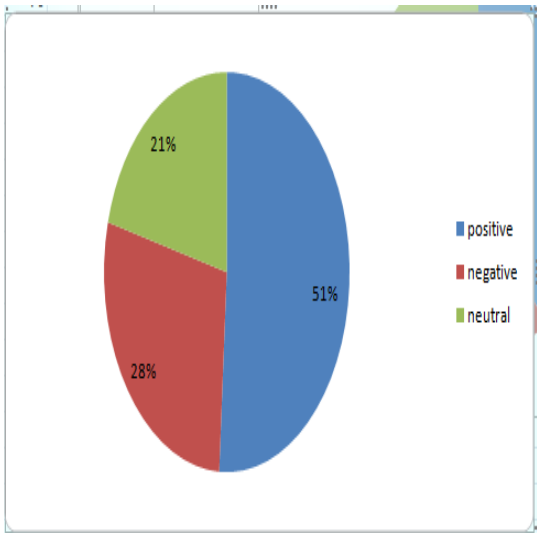

Figure 11: sentiment of text written in English language

In machine learning algorithm we have used three classification based algorithm namely Perceptron, Bayes, and Rocchio. In arrange way to evaluate their performance in predicting whether sentiment of cleaned text file of Facebook is positive or negative we have used machine learning algorithm only for English written text. Second table shows the important metrics and performance of the classifiers. we have collected comments of 4079 posts of Facebook publicly available page into the one text file. We have taken the following allotment of dataset for training and testing sets which is shown in table 1

Table 1: Data Distribution

Sentiment	Training	Testing
Positive	2000	2000
Negative	2000	2000

The three classifiers were compared in term of these three metrics: Precision, Recall and F-Score performance using the computation shown below

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

$$F - Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

	Naïve Bayes Classifier	Perceptron Classifier	Rocchio Classifier
Precision	94%	78%	92%
Recall	96%	76%	89%
F-Score	93%	74%	87%

Table 2: Metrics of three classifiers for Bigram of BOW feature

	Naïve Bayes Classifier	Perceptron Classifier	Rocchio Classifier
Precision	79%	72%	74%
Recall	75%	73%	76%
F-Score	76%	71%	73%

Table 3: Metrics of three classifiers for Trigram of BOW feature

If we compare metrics of table 2 with table 3, all entries of table 2 are high compared to the table 3. So our assorted sentiment analyzer performed well for bigram of BOW feature.

V. CONCLUSION

we have implemented assorted sentiment analyzer model for publically available page of facebook which is a worldwide famous social media site. We have taken publically available page of prime minister of India (narendra modi) from facebook for checking the performance of our sentiment analyzer model. It works fine for Bigram for BOW feature of machine learning algorithm compared to trigram feature of BOW. Our Data extractor model fetch comments of all the post of page one text file in the less span of time compared to API Time complexity of our assorted sentiment analysis model in the best case is O (logn). Best case is found after applying CILTEL model for finding sentiment of publically available page of Facebook and in the worst case it is O (N). Worst case is found without CILTEL model for finding sentiment of publically available page of Facebook. Our model is beneficial for industry to check the user prospective about the posts related to product of company in its publically available page of facebook. Similarly it is also beneficially for publically available page of political parties and leader after checking the sentiments of the user

REFERENCES

1. A.M. Kaplan, M. Heinlein, Users of the world, unite! The challenges and opportunities of Social Media, Business Horizons 53 (1) (2010) 59–68.
2. B. Pang, L. Lee, Opinion Mining and Sentiment Analysis, Foundations and Trends in Information Retrieval 2. No. 2 in 2., Now Publishers Inc. 2008, pp. 1–135.
3. Liu. B, Synthesis Lectures on Human Language Technologies, Sentiment Analysis and Opinion Mining 5, Morgan & Claypool Publishers. 2012, pp. 1–167
4. J. Scheibman, Point of View and Grammar: Structural Patterns of Subjectivity in American English Conversation, vol. 11, John Benjamin Publishing, 2002.
5. W. Langacker, Observations and speculations on subjectivity, Iconicity Syntax 1 (985) (1985) 109.
6. A. Agarwal, B. Xie, I. Vovsha, O. Rambow, R. Passonneau, "Sentiment analysis of Twitter data", LSM '11 Proceedings of the Workshop on Languages in Social Media, Association for Computational Linguistics, pp. 30-38, 2011.
7. B.J. Jansen, M. Zhang, K.Sobel,A. Chowdury, twitter power: tweets as electronic word of mouth, J.AM. Society of information science 60(11) 2009 2169-2188
8. O. Araque, I. Corcuera-Platas, C.A. Iglesias, Enhancing deep learning sentiment analysis with ensemble techniques in social applications, Expert System Application 2017
9. I. Chaturvedi, E. Cambria, F. Herrera, Distinguishing between facts and opinions for sentiment analysis: survey and challenges, Inf. Fusion 44(2018) 65-77.
10. K. Ravi, V. Ravi, A Survey on opinion mining and sentiment analysis: tasks, approaches and applications, Knowledge Based System 89 (2015) 14-46
11. Y. Genc, Y. Sakamoto, J. Nickerson, Discovering context: classifying tweets through a semantic transform based on Wikipedia, foundations of augmented cognition, directing the future of adaptive systems (2011) 484-492
12. W. Guo, H. Li, M. T. Diab, linking tweets to news: a framework to enrich short text data in social media, in: ACL (1) 2013 239-249
13. S. Poria, E.Cambria, R.Bajpai, A.Hussain, A Review of affective computing: from unimodal analysis to multimodal fusion, information fusion 37(2017) 98-125
14. S. Volkora, predicting demographics and affect in social networks johns Hopkins university Baltimore Maryland 2015 thesis
15. D.Hovy, Demographic factors improve classification performance in: ACL(1) 2015 pp 752-762
16. H.Chen, J. Liu, M.Liu, Q.Zheng , Semi-supervised clue fusion for spammer detection in Sina Weibo, information fusion 44(2018) 22-32
17. Inside Facebook: http://www.insidefacebook.com/2011/09/21/5000-character-limit-floating-navigation-bar/
18. Liu.B. , sentiment analysis and subjectivity, Handbook of natural language processing pp 627-666
19. S.Wang, C.D.Manning Baselines and bigrams: simple, good sentiment and topic classification, in: association for computer linguistics stroudsberg USA 2012 pp. 90-94
20. A.Sharma, S.Dey, A Comparative Study of feature selection and machine learning techniques for sentiment analysis, in: proceeding of the 2012 ACM Research in Applied Computation Symposium ACM, 2012, pp 1-7
21. B. Pang, L. Lee, S. Vaithyanathan , Thumb Up? : Sentiment classification using machine learning techniques in: proceedings of empirical methods in natural language processing volume 10 pp 79-86
22. Das. S &Chen.M , Yahoo for Amazon: Extracting market sentiment from stock message boards . In Proceedings of the Asia Pacific Finance association annual conference 2001
23. Spertus, Smokey: Automatic recognition of hostile messages, in proceedings of innovative application of artificial intelligence 1997
24. Mishne & Glance, Predicting movie sales from blogger sentiment, in AAAI Symposium on computational approaches to analysing weblogs 2006 pp 155-158
25. Feldman.R, Techniques and application for sentiment analysis, communication of the ACM, 2013, 56(4) pp 82-89
26. Abbasi, Affect Intensity analysis of dark web forums, in proceedings of intelligence and security informatics 2007, pp 282-288

Retrieval Number: B7739129219/2020©BEIESP
DOI: 10.35940/ijitee.B7739.019320
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

27. Carro, Pulido,Rodriguez , Dynamic generation of adaptive based courses, Journal of Networks and Computer Application 1999, 22, pp 249-257
28. Liu. H., Liberman.H & Selker.T., A model of textual affect sensing using real word knowledge, proceedings of intelligent user interfaces 2003, pp 125-132
29. Rogers(2003), Diffusion of innovation, New York: Free Press
30. Turney(2002) Thumbs up or Thumbs down? : Semantic orientation applied to unsupervised classification of reviews, in proceedings of association for computational linguistics 2002, pp 417-424
31. B. Pang, L. Lee, “A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts”, ACL '04 Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics, pp. 271-278, 2004
32. E. Boiy, P. Hens, K. Deschacht, M. - F. Moens, “Automatic Sentiment Analysis in On-line Text”, Proceedings of the 11th International Conference on Electronic Publishing, pp. 349-360, 2007.
33. L. Barbosa, J. Feng, “Robust sentiment detection on twitter from biased and noisy data”, Proceedings of the 23rd International Conference on Computational Linguistics, pp. 36–44, 2010
34. M. Gamon, “Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis”, Proceedings of the 20th international conference on Computational Linguistics, pp. 841–847, 2004.
35. Fang, X., & Zhan, J. (2015). Sentiment analysis using product review data. Journal of Big Data, 2(5), 1–14. http://dx.doi.org/10.1186/s40537-015-0015-2
36. Jeong, B., Yoon, J., & Lee, J. M. (2017). Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis. International Journal of Information Management. http://dx.doi.org/10.1016/j.ijinfomgt.2017.09.009 (In Press)
37. Hu, Y. H., & Chen, K. (2016). Predicting hotel review helpfulness: The impact of review visibility, and interaction between hotel stars and review ratings. International Journal of Information Management, 36(6), 929–944
38. He, W., Zha, S., & Li, L. (2013). Social media competitive analysis and text mining: A case study in the pizza industry. International Journal of Information Management, 33(3), 464–472.
39. Dave, Lawrence & Pennock, mining the peanut gallery: opinion extraction and semantic classification of product reviews in proceeding of world wide web 2003 pp 519-528

AUTHORS PROFILE

Saurabh Dhyani received Graduation degree from Hemwati nandan bahuguna Garhwal University , and Post Graduation Degree in Computer Applications from University of Rajasthan. He is currently Pursuing PhD Degree in the department of Computer Applications, Maulana Azad National Institute of technology. He has three year teaching experience. His research interests include Sentiment Analysis, Text Mining, Web Usage Mining and Social Networking.

Ghanshyam Singh Thakur has received BSc degree from Dr. Hari Singh Gour University Sagar M.P. in 2000. He has received MCA degree in 2003 from Pt. RaviShankar Shukal University Raipur C.G. and PhD degree from Barkhatullah University, Bhopal M.P. in year 2009. He is Assistant Professor in the department of Computer Applications, Maulana Azad National Institute of technology, Bhopal, M. P. India. He has eight year teaching and research experience. He has 26 publications in national and international journals. His research interests include Text Mining, Document clustering, Information Retrieval, Data Warehousing. He is a member of the CSI, IAENG, and IACSIT.