Flirt Analysis In Private Chat Using Natural Language Processing

[1]V.Selina Annie Retna, [2] Prof. P. Brundha

[1] PG Student, Francis Xavier Engineering College, Department of CSE

[2]Head of the Department, Francis Xavier Engineering College, Department of CSE

ABSTRACT

Every mode of communication now became through online and everyone prefers the online communication than all others. Considering the business people who runs the industry either a large sector or the normal middle sized industry every mode of communication and reaching for their product now started reaching through online mode. By coming through the communication everyone most likely prefers through WhatsApp communication. Not only had the business people also every normal people preferred to communicate with one another through common chat application. By using this mode everyone will come to a state how our chat was going and to analyse the conversation was going, for that it was using the flirt analysis using sentimental analysis. The flirt analysis is done by the mining concept and the exact result by natural language processing.

Keywords: Sentimental Analysis, Data mining, Emotional, Natural Language Processing

Article Received: 18 October 2020, Revised: 3 November 2020, Accepted: 24 December 2020

INTRODUCTION

All sort of communication and also the usage of the social media nowadays bloom drastically it was the easiest and the fastest means of communication and also the easiest way to connect with one another. Not only the easiest way to connect with one another but also to check the latest update regarding the product and the review of the product can also be reached through online which was very helpful in order to reach out the best product and also connected with the known group of peoples. There comes a question whether all the conversation and the process we are making for the search are going correct and also the analyze our own product review and based upon which every product can be changed and equipped by the valid review given by the customers. By considering with the personal set of enquires whether all our chat with the person was going good and also the person who wants to monitor with their children who was using their chat communication whether all their chat was going in a correct manner or they are stepping out the way we are into.

LITERATURE SURVEY

Sentimental part was mostly connected to our day to day life and was connected with our self which we need and expect in our life natural events. Everyone like to get feedback for what are all the things we are doing and was ready to change if they are going in the wrong way. So in this technological life to detect about the sentimental detection it will deals with some standard method of detecting the textual method by the sentimental analysis algorithms.

Considering with the brand recognizes and the exact reviews and the follow up thus deal with the regular substructures in the method of detecting the sentimental part from the customers. Several algorithms were present in order which we can find out the sentimental reviles of the product of

the bran and also the reach to the exact number of peoples.

Based upon the need of analyses and also how hard the process of analyzing part was it was detected for which mode of algorithm was used for the mining of the textual concept of the conclusion. Each algorithm comprises of certain structure and will be used best with their correct mode of utilization.

While considering the twitter analysis, the data which was given in the twitter handle will be analyzed for the sentimental review process and also regarding the analysis the exact result was received as the result was obtained by using the relevant algorithm.

In Facebook the sentimental analysis is based on the public opinion from the peoples who can be a customer, user or the person who experienced the brand or the product which was made for the usage of the people. From the polling or the review taken from the public point, the view will be used for understanding how the product is and based upon which they can also brand them according to their reviews they received.

The sentimental analysis in the education sector was used for the analyzation of the fact to the students whether they understand the concept or how about their understanding towards the subject was analyzed. For this type of analysis the chat bot concept was introduced from that it was nowadays excelled and based upon the sentimental analysis it was given out the required result. Based upon the current mood of the person the surroundings were modified to make chill. For this entire concept the sentimental analysis was used.

Coming to the analysis part in the email processing few of the mail which we received will be automatically transferred to the spam or junk folder, the reason behind the process was it will be automatically detecting the messages with the standard set of list which will be checked with the message if certain criteria was matched with the received mail it will automatically process and move them to the junk folder.

Mining concept was now a day's everywhere for the analysis of the text or else any conceptual structures and reaching out with the exact results. While considering the flirt analysis the text mining concept was most commonly used for the analysis as it was the concept which will be compared and the exact data will be compared with the given set of details and the respective results was produced.

PROPOSED SYSTEM

The sentimental analysis was the widely used analysis method in order to find out the feedback or the review among the peoples and to give out the exact details regarding the results.

Flirt analysis is the algorithm which was implemented in order to identify the flirt concept in the chat message to detect and analyze the chat messages and to compare and group the set of values related with the given set if values in the flirt list to identify the most flirted conversation

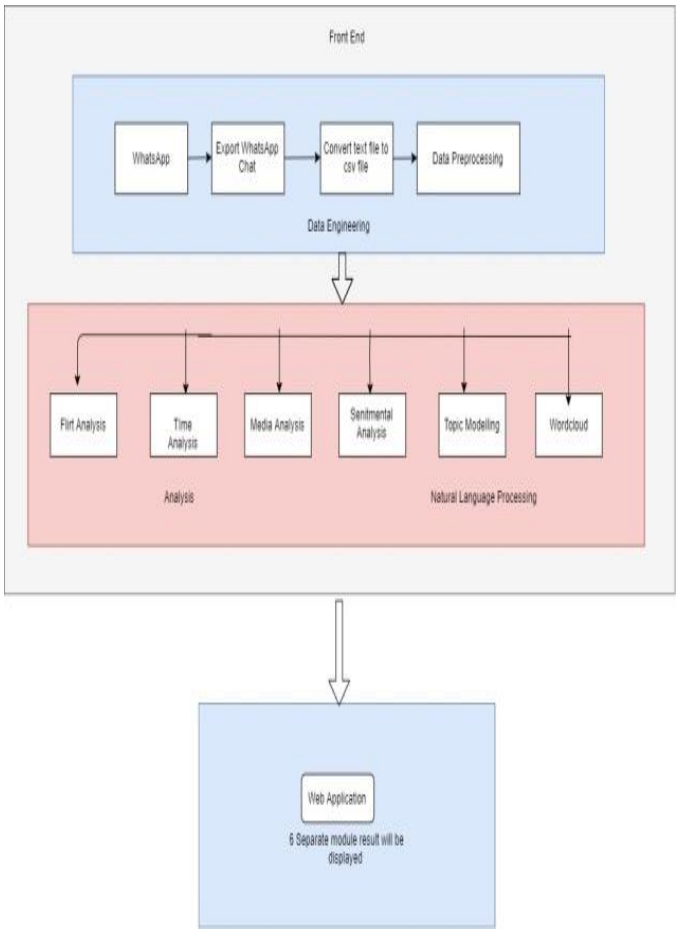

Fig 1. Flirt Analysis Module

Data Collection:

While processing or want to get any result we need some data to be get processed to obtain the result. In the flirt analysis concept the input data and the process which will be taken place in the front end will be collecting the data which was the chat details and this can be done by exporting the chat from the chat application and that was taken out for the input (exported chat). The chat was converted into csv file for further process.

Data Preprocessing:

Every set of input will be needed to perform certain steps of preprocessing and it was done for the easy evaluation and also for the better work flow regarding every step in the process.

Hence the preprocessing step was the mandatory process in the executing the process. The steps for the preprocessing were listed below:

1. The excessive space in the chat data was first reduced for the process
2. The emoji which was used in the chat was selected and then it was subjected to the ignorance part
3. The date wise chat arrangement was made which will be easy for the processing of the data which was given as the input.

Flirt Analysis:

The analyzing of the given set of data which was given as the input was considered and compared to the respective set of flirt data set and the result will be displayed according to the compared result.

Flirt Analysis Algorithm

Step 1: The set of flirt words was selected and was feed into the system.

Step 2: Group the message according to the respective sender of the chat.

Step 3: Calculating the uniqueness of the conversation in the chat (U)

Step 4: the total number of words which was used in each row was analysed and was given out (T_w)

Step 5: The repetition of the word was calculated and was identified (U_w)

Step 6: Formula: Unique Frequency for a Word (U_w) = (U/sum of T_w)* 100

Step 7: The total number of the flirt words which matches in the list was calculated (F_l)

Step 8: Check each word with a collected flirt list

Step 9: The flirt word used by each person was identified and grouped. (FW)

Step 10: Calculate the total number of flirt words used by each person. (L_f)

Step 11: Unique flirt words used by the respective person were identified (F_f)

Step 12: Formula for Unique Flirt frequency (F_f) = (U/F_l)*100

Step 13: Calculate Total Flirt Percentage=sum of F_F/ total number of F_w.

Time Analysis:

Time analysis is the grouping of the chat according to the respective time was made and was projected out for the analysing process and then they were subjected for the grouping them

Max (Repeated date, Date, Time)

Media Analysis:

In this analysis the media was taken for example the picture the voice note or the video or call which was shared between the two people was grouped and was analyzed and then taken out for the consideration and was grouped for the analysis part and the result was made.

Media shared count [Query (Select the media shared)for each person in the chat]

Topic Modelling:

Based upon which the required set of the chat which we are gone into was made for the analysis part and thus was given the respective topic for the chat which was processed and was then group under the respective topic and was given out as the expected results.

Word Cloud:

The word cloud is the visual representation of the word which was used in the conversation. The frequently used word will be taken the place of the bigger font size and respective the least used will taken up the small font size

Django Framework

Django framework is the web development framework. By this we can implement all the module result into the structured values which will be displayed in the web page by using the local host.

Performance Analysis:

1. The project was made to undergone the sentimental flirt analysis in the private setting of a group in the WhatsApp chat and it was implemented and was given out the exact result for the analysis
2. The working process was undergone with first starting with the pre processing steps and after that, it basically checks the input with the flirt list and thus processes and then the result was to be displayed. Exactly what was the expectation from this flirt analysis methodology was obtained in this analyzing process
3. The procedures were created so that they can logically perform the sub-functions for the flirt analysis technology
4. All the codes which were given for obtaining the result were readable and it can be easily taken for the understanding

and the structure of the progress all was implemented according to the procedures.

PSEUDOCODE

Inputs:

$S=[s_1,\ldots,sn\,J, s_1 E[-100,100)$ tweet'ssentimen tvector

$E=[e_1,\ldots,en], e_1 E[O, 1)$ tweet's evidencevector

$p=4.5$

ClassificationProcess:

$Wp=$ i

IF$(WN==0)$

RETURN Fp

ELSEIF$(WP==0)$

RETURN FN

ELSE{

$Ep=$ Lei>0.55s1>01

}

RETURNFp

ELSEIF$(FN-Fp>0.1)$

RETURNFN

ELSE{

IF$(Fp+FN>0)$

RETURNFp

ELSEIF$(Fp+FN<0)$

RETURNFN

ELSE

RETURN0

}

RESULT:

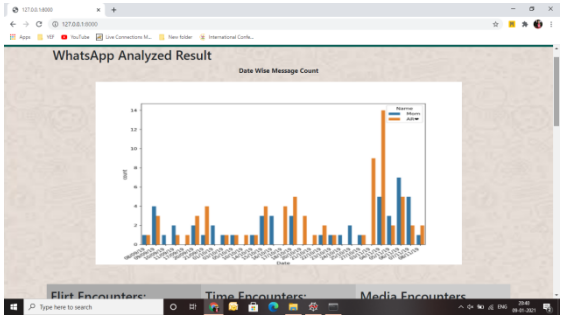

Fig2. Date wise analysis

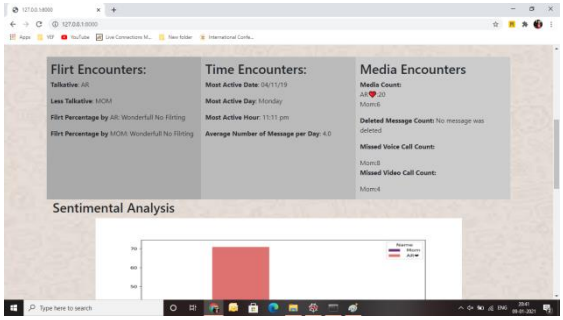

Fig3. Time and media encounters

Fig4. Topic Modelling

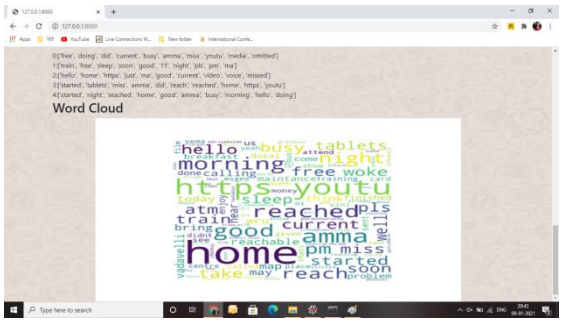

Fig5. Word Cloud

CONCLUSION AND FUTUREWORK

In this flirt analysis concept the personal chat was taken in and was processed for obtaining the flirt percentage in the given set of chat and was resulted with the accurate process.

The security methods can be included in the future upcoming implementation part and also the values of maintaining with the chat data will cost considering that will be effective.

REFERENCES

[1]"Sentiment analysis of Twitter data during critical events through Bayesian networks classifiers",Gonzalo A.Ruz[ab]Pablo A.Henríquez[a]AldoMascareño-8 January 2020

[2] Mohammad A.Hassonah, RizikAl-Sayyed, AliRodan, Ala' M.Al-Zoubi[c]IbrahimAljarah[a]HossamFaris, "An efficient hybrid filter and evolutionary wrapper approach for sentiment analysis of various topics on Twitter"- 11 December 2019

[3]Shikha Tiwari; Anshika Verma; Peeyush Garg; Deepika Bansal,"Social Media Sentiment Analysis On Twitter Datasets"

23 April 2020 -**ISSN:** 2575-7288

"Sentiment analysis using deep learning architectures: a review" Ashima Yadav &

Dinesh Kumar Vishwakarma- 2 dec- 2020

[4] B. Liu and L. Zhang, "A survey of opinion mining and sentiment analysis," in mining text data: Springer, 2012, pp. 415-463.

[5] Y. Wang, "Sensing Human Sentiment via Social Media Images: Methodologies and Applications," Arizona State University, 2018.

[6]S. D. Pressman, M. W. Gallagher, and S. J. Lopez, "Is the emotion-health connection a "first-world problem"?," Psychological science, vol. 24, no. 4, pp. 544-549, 2013

[7]P. Garg, H. Garg, and V. Ranga, "Sentiment analysis of the Uri terror attack using Twitter," in 2017 International Conference on Computing, Communication and Automation (ICCCA), 2017: IEEE, pp. 17-20.

[8] B. O'Connor, R. Balasubramanyan, B. R. Routledge, and N. A. Smith, "From tweets to polls: Linking text sentiment to public opinion time series.," Icwsm, vol. 11, no. 122–129, pp. 1–2, 2010.

[9] M. A. Cabanlit and K. J. Espinosa, "Optimizing N-gram based text feature selection in sentiment analysis for commercial products in Twitter through polarity lexicons," in Information, Intelligence, Systems and Applications, IISA 2014, The 5th International Conference on, 2014, pp. 94–97.

[10] S.-M. Kim and E. Hovy, "Determining the sentiment of opinions," in Proceedings of the 20th international conference on Computational Linguistics, 2004, p. 1367.

[11] C. Whitelaw, N. Garg, and S. Argamon, "Using appraisal groups for sentiment analysis," in Proceedings of the 14th ACM international conference on Information and knowledge management, 2005, pp. 625–631.

[12] H. Saif, M. Fernandez, Y. He, and H. Alani, "Evaluation datasets for Twitter sentiment analysis," Emot. Sentiment. Soc. Expressive Media, p. 9, 2013.

[13] A. P. Jain and P. Dandannavar, "Application of machine learning techniques to sentiment analysis," in 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), 2016, pp. 628–632.

[14] A. Go, R. Bhayani, and L. Huang, "Twitter Sentiment Classification using Distant Supervision," Processing, vol. 150, no. 12, pp. 1–6, 2009.

[15] V. Singh and S. K. Dubey, "Opinion mining and analysis: A literature review," in 2014 5th

International Conference-Confluence The Next Generation Information Technology Summit (Confluence), 2014, pp. 232–239.

[16] W. Chu and S. S. Keerthi, "Support vector ordinal regression," Neural Comput., vol. 19, no. 3, pp. 792–815, 2007.

[17] S. Liu, F. Li, F. Li, X. Cheng, and H. Shen, "Adaptive co-training SVM for sentiment classification on tweets," in Proceedings of the 22nd ACM international conference on Information & Knowledge Management, 2013, pp. 2079–2088

[18] B. Liu and L. Zhang, "A survey of opinion mining and sentiment analysis," in Mining text data: Springer, 2012, pp. 415-463.

[19] Manikandan, R and Dr.R.Latha (2017). "A literature survey of existing map matching algorithm for navigation technology. International journal of engineering sciences & research technology", 6(9), 326-331.Retrieved September 15, 2017.

[20] A.M. Barani, R.Latha, R.Manikandan, "Implementation of Artificial Fish Swarm Optimization for Cardiovascular Heart Disease" International Journal of Recent Technology and Engineering (IJRTE), Vol. 08, No. 4S5, 134-136, 2019.

[21] Manikandan, R., Latha, R., & Ambethraj, C. (1). An Analysis of Map Matching Algorithm for Recent Intelligent Transport System. Asian Journal of Applied Sciences, 5(1). Retrieved from https://www.ajouronline.com/index.php/AJAS/article/view/4642.

[22] R. Sathish, R. Manikandan, S. Silvia Priscila, B. V. Sara and R. Mahaveerakannan, "A Report on the Impact of Information Technology and Social Media on Covid–19," 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), Thoothukudi, India, 2020, pp. 224-230, doi: 10.1109/ICISS49785.2020.9316046.

[23] Manikandan, R and Dr.R.Latha (2018). "Map Matching Algorithm Based on a Hidden Markov Model for Vehicle Navigation" International Journal of Advanced Technology in Engineering and Science, 6(6), 36-42.

[24] Manikandan, R and Dr.R.Latha (2018). "GLOBAL POSITIONING SYSTEM FOR VEHICLE NAVIGATION" International Journal of Advances in Arts, Sciences and Engineering (IJOAASE), 6(13), 1-9.