Employee Interaction Analysis In Cooperate Environment

Gangatharan S, Lakshmi N, Vishnupriya B, Vijayakumar D

Abstract *--To develop an environment in which the organization can analyze the employee's technical skills and interest towards the projects. Social media has received more attention nowadays. Public and private opinion about a wide variety of subjects is expressed and spread continually via many social media. Based on the analysis report the organization head will distribute the project for the correct person. Developing a program employee interaction analysis is an approach to be used to computationally measure employee' perceptions. In this paper, employees are combined in to a project group chat. Our paper aims to analyze these chats and to classify them as positive, weakly positive, strongly positive, weakly negative, strongly negative, negative and for neutral based on polarity values of chats. The aim of this paper is to analysis the pros and cons of the employee. This will help the organization to find the suitable person for the required project. The system finds this worth team through visualized network graph.*

Index Terms *- Network, NLP, Character, Co-occurrence, PySpark*

I. INTRODUCTION

Now, the internet is becoming more efficient as horizon become wider. Social media like Twitter, Blogging are more efficient in spreading news across the world. A topic become trending if more users are contributing their opinion towards it. For an instance, they leak the images of the new upcoming product. From the social media, they gather the emotions of the people towards that product. Employee interaction analysis is the prediction of emotions in a word or in a sentence. It is serve as an application to understand the attitudes, emotions in an online application. It is the process of categorizing conversation into positive, negative or neutral labels. Normally, people use social media sites for interacting with other people and to stay up-to-date with news and current events. For an instance, people quickly share the reviews on online as soon as they watched the movie and the interaction continued as series of comments to discuss about the movie. This type of information on these sites is used for marketing. Therefore, Employee Interaction analysis includes emotion mining, polarity and classification. Employee Interaction Analysis involves the use of Natural Language

Processing to extract, identify to characterize the sentiment. A text may include two types of sentiments Direct and Comparative. In direct sentiments, the object of the sentence is independent of each other. In comparative sentiments, they compare the object of the same sentence. Employee Interaction analysis is the process of analyzing the efficient team for the project. Here, the system takes the online chat as input. The output depicts the most interactive persons in a team and most interactive team among all the teams.

II. LITERARTURY SURVEY

Most of the world's data is unstructured, in other words it is unorganized. Huge amount of data is created day-to-day, which is hard to analyze, sort through. In addition to it also cost expensive and time consuming. It is one of the most important thing in understanding people emotion for the project, that is their own individual thought about the project. By analyzing history of chats, we can able to understand the behavior and thought of a team towards that project.

It is difficult for us to manage the large data and sort through it manually. Sentimental analysis helps in sorting those data and produce the result accordingly. By using sentimental analysis, we an able to obtain only 60 to 65 percent of accuracy in our result. So, that organization prefer centralized sentiment analysis system, it helps in improve accuracy. Understanding people emotions is essential for an organization as they can able to predict the mindset of an employee related to that project. For an instance, consider an online shopping like amazon, in that the customer post the comments for the project. So, that anyone on this site can able to know about the pros and the cons of the product. The customer can aware of the cons of the product. There may fluctuating in the comments. Some customer can't able to predict the product. So, the organization in need to clarify the customer opinion. So, they do the sentimental analysis of the customer reviews. From this analysis, they able to classify the review of the customer and predict their opinion as positive, negative or neutral.

Opinion and sentiment mining is an important research area because due to the huge number of daily posts on social networks, extracting people's opinion is a challenging task.

III. MATERIALS ANDMETHODS

Requirements

Operating System: Linux

Data Extraction : WhatsApp account Data

Extracted by WhatsApp chat cleaner

Data Identification: NLTK, PySpark, NumPy

GUI: Python frame work Change to table format

In the proposed system, the system analyzes both the emotions and the technical knowledge of the persons by Natural Language Processing. It uses python at the backend. Python is a high-level, interpreted and general- purpose dynamic programming language that focuses on code readability. The syntax in Python helps the programmers to do coding in fewer steps as compared to Java. The language founded in the year 1991 by the developer Guido Van Rossum made the programming easy and fun to do. The Python is widely used in bigger organizations because of its multiple programming paradigms. They usually involve imperative and object- oriented functional programming. It has a comprehensive and large standard library that has automatic memory management and dynamic features. In sentimental analysis, normally we analyze the emotions of the persons and predict the output. Normally, in sentimental analysis we analyze the emotions of the persons and produce the overall output. For an instance, consider sentimental analysis in twitter, here we analyze entire tweets and predict the output as emotions of a people related to a product or event.

System Architecture

Fig 3.1 represents the architecture diagram of the project we take one chatting environment and post our project and allow all the teams to interact and analyze the team interaction and their technical knowledge. Based on the analyses report, the system selects the teams and assign project to them. This is to analyze which team can produce the efficient output for that project. This proposed system helped the organization to find out the right team for the project. Normally, in earlier days in the organization they conduct the interview to pick the persons who suits for the project. This may take long time to find out efficient team and we are in need some human power to find out the efficient team. With the help of the proposed system organization can easily identify the efficient team. They produce the output as the visualized network graph. The above proposed system can be analyzed using the natural language processing. Natural language processing (NLP) describes the interaction between human language and computers. ... A few examples of NLP that people use every day are: Spell check. Spell check is a software program that corrects spelling errors in word processing, email and online discussions. Spell check identifies and corrects misspelled words. ... In Microsoft Word, spell check options, like spelling and grammar may be found

under the 'review' tab and 'proofing' window. In our proposed system we produce the output as the file. It contains the names of the persons, who present in the team. We can also predict this by probability. **3.3 3.33.3**

Equations

Naive Bayes Classifier: It is used to predict the probability for given words that belong to a cleans the unstructured textual data into a structured textual class. It is used because of its easiness in both training and classifying steps. Pre-processed data is given as input to train input set using Naive Bayes classifier and that trained model is applied on the test to generate the sentiment.
The Bayes theorem is asfollows.

$P(H|X) = P(X|H).P(H)$ ----- eq(1)
P(X) represents X- Tuples H represents Hypothesis
P(H|X) represents Posterior probability of H conditioned on X the Probability that Hypothesis holds true given the value of X
P(H) represents Prior probability of H (i.e.) the Probability that H holds true irrespective of the tuple values
P(X|H) represents posterior probability of X conditioned on H i.e. tested based on Chi-square method. It creates a list of all the Probability that X will have certain values for a given positive and negative words.

P(X) represents Prior probability of X i.e. the distribution Probability that X will have certain values. An experiment result of accuracy is evaluated using following information retrieval matrices. Accuracy is the result for both positive and negative classes and it is a performance evaluation parameter and it is calculated by the number of correctly selected positive and negative words word scores are found and the best number of words divided by the total number of words present in the corpus.

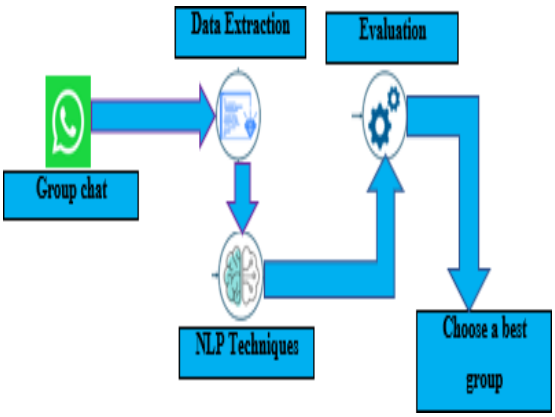
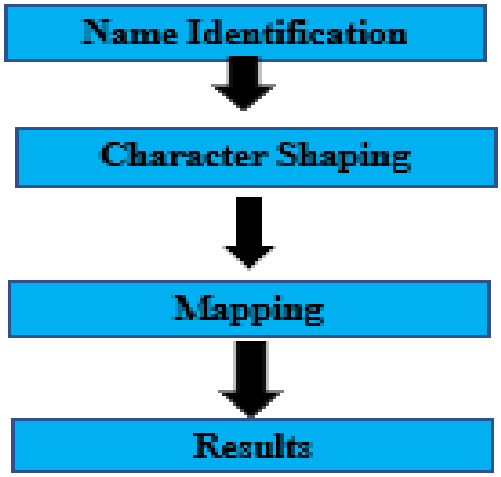


Fig 3.1 Architecture Diagram



Data flow diagram

Name identification
Identifies an employee names in each group.

Character Shaping

Create a file that has the project related technical terms, and then we can compare it with the employee's conversation. By using this technique, we can gather the behavior of the employees

Mapping

Based on the character shaping we can map their interactions with colleagues to get the exact map of them.

Result
At last the decision is made to select the right team for the project according to the mapping.

IV. RESULT AND DISCUSSIONS

The final output generated by the system is in the form of graphs and the file which is generated along with it. In that graph the relationship between the two nodes (i.e. Two Characters) are clearly explained by the thickness of colors. The thicker the color is the closer the relationship is. Fig 4.1 The Co-occurrence graph clearly explains the connection between the nodes and the depth of the connections is explained in sentiment graph (Fig 4.2). The values of the nodes are written in the file for further reference.

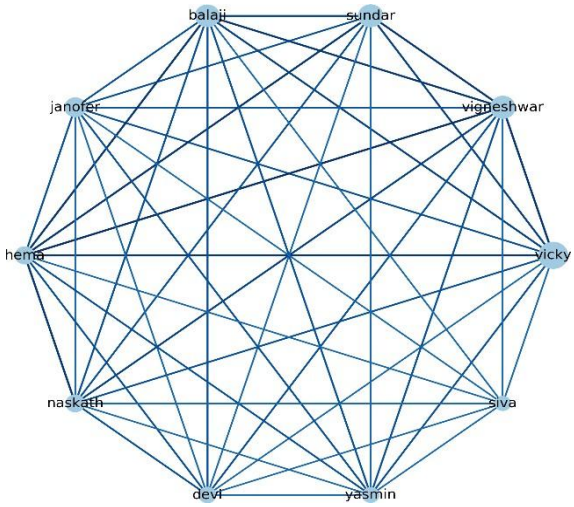


Fig4.1 co-occurrence Graph

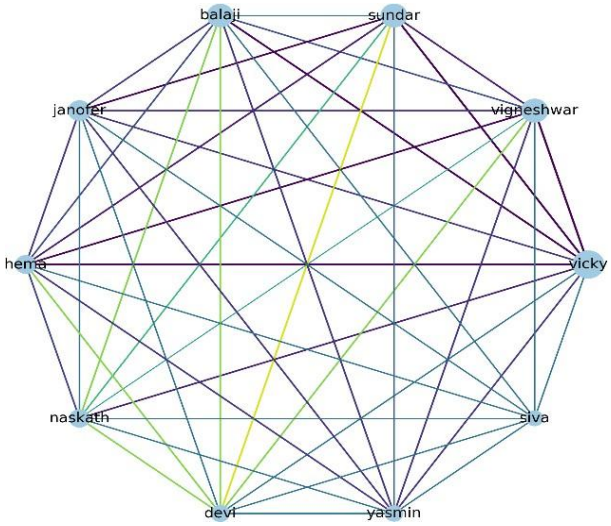


Fig4.2 Sentiment Graph

V. PREPARATION OF TABLES

Weights

4.457 5	4.613 4	4.129 2	3.973 5	4.256 5	3.772	4.129 2
4.055 5	4.412 4	4.412 4	3.973 5	3.973 5	4.055 6	3.772 9
4.256 5	5.320 9	4.129 2	3.880 3	4.129 2	3.290 9	4.412 4
4.129 2	3.880 3	4.129 2	3.290 9	4.412 4	4.129 2	3.772 9
3.010 2	3.973 5	3.645 9	3.880 3	3.880 3	3.772 9	3.645 9
4.055 6	3.973 5	2.534 4	3.290 9	4.055 6	3.490 7	4.756 4

Colors

6.797 5	6.873 4	6.439 2	5.173 5	5.256 5	3.772 1	7.129 2
6.875 6,	6.712 4	6.902 4	5.123 5	4.973 5	5.055 6	8.772 9
5.886 5	6.567 9	8.899 2	7.650 3	5.129 2	6.290 9	8.412 4
4.769 2	7.880 3	6.679 2	4.430 9	7.412 4	8.129 2	9.772 9
6.650 2	7.867 1	7.231 5	9.430 3	7.880 3	8.772 9	5.645 9
7.785 6	5.773 5	9.344 4	4.670 9	9.055 6	9.490 7	7.756 4

VI. CONCLUSION

Employee analysis comprises of steps like data collection, sentimental detection, comparing it with our threshold value and predicts the result. The company or corporation analyzes employee interaction via group chat. Using this tool, the company easily identifies the employee's technical knowledge and apt the best group for the project. The project will be completed within a short time using this tool because this tool identifies the best group for this project. So, the best team will perform completed the project very quickly and efficiently. Also, this tool will help the employee to find their strength and weakness in technical aspects.

VII. REFERENCES

[1] Directed Graphs, Multigraphs and Visualization in Network x byArijit Gayen
[2] Sentiment Analysis with AFINN Lexicon by Himanshu Lohiya(jan-30)
[3] Tokenizing Words and Sentences with NLTK by python.org
[4] Sentiment Analysis Basics bypython.org
[5] Sentiment Analysis: Types,Tools, and Use Cases by Altexsoft(sep-2018)