An Employee Feedback Model Based on Sentimental Analysis and Word Clouds

Janhvi Juyal[1], Sonakshi Vij[2], Deepali Virmani[3]

Vivekananda Institute of Professional Studies – Technical Campus, GGSIPU

Corresponding author: Janhvi Juyal, Email: Juyaljanhvi09@gmail.com

Employee feedback is significant for any organization to perceive what's going on in the minds of every employee. It increases the engagement, efficiency, and productivity of employees and makes them feel valued. Both positive and negative feedbacks are beneficial for any organization. With the help of positive feedback, it is known what the employee likes about the company and what is the strong part of it and how can it become better. It increases the overall gain of the company. Negative feedback should be taken as constructive criticism i.e., the organization should work where it's lagging to fulfil the wishes and needs of employees. This paper proposes a method for employee feedback analysis using sentiment analysis through the generation of word clouds. To test this feedback model TCS Employee review data was taken into consideration which consisted of Indian as well as International Reviews. The results show a positive inclination towards the company and represent that people are generally happy and satisfied with the organization.

Keywords: Sentiment Analysis, Feedback, Employee Reviews, Review Analysis, TCS, Word Cloud

Janhvi Juyal[1], Sonakshi Vij[2], Deepali Virmani[3]

1. Introduction

To develop any organization there is always a need for constructive criticism. Constructive criticism is the feedback given by the consumers or the costumers for the product, publication, services, etc. provided by the organization. This constructive criticism helps the organization to grow professionally and helps to develop strong professional relationships. These feedbacks are known as Reviews. The reviews give an organization a clear picture of its strengths and weaknesses and help to develop its long- and short-term goals according to that [1-8]. Broadly reviews are divided into two categories: Positive and Negative. Positive reviews are generally in the favor of the organization. It helps the organization to develop goodwill and a positive reputation in the market. It affects the website's ranking positively. The communication process is improved by the reviews as the organization is aware of what their employees need and it helps them to understand the mindset of their employees [17-18]. On the other hand, negative reviews are a reflection of the unsatisfactory services provided by the organization to the employees or consumers. It helps the organization to identify in which field they are lagging and how they can improve their services and expand their business. These reviews bring worthy feedback and potentially help to enhance the organization by providing insights into consumers' behavior and sentiments [13-14].

Reviews are collected from social media or customer feedback forms or online surveys etc. The collection of these reviews is the first step toward sentimental analysis [3]. Sentimental analysis is the methodology to gauge or analyze the emotion depicted by the reviews. Machine learning, Natural Language Processing (NLP), and statistics are some basic techniques used in sentimental analysis to scratch out the feelings and thoughts that are conveyed by people. Few algorithms that are conventionally used by the sentimental analysis model are mentioned below:

1. Hybrid: This particular method is done manually and automatically. It is the most precise and accurate approach.
2. Automatic or Machine Learning based analysis: Datasets that consist of positive or negative words are used to train the Machine learning models [6]. This model can detect or differentiate between the positive or negative words contained in the review once the training is complete.
3. Rule-Based Analysis: The polarity of the words present in the reviews is the major aspect of this analysis. It counts the number of negative and positive words and depicts the sentiment or tone of the text. Lexicons – a list of words and sentiments that are imparted by them – are the type of NLP technique that helps the system to identify the sentiments by counting the polarized words [4].

In the paper, the author referred to the method of Ruled Based Analysis for the sentimental analysis of the feedback or reviews given by the employees. This method provided an accurate direction to the analysis. The proposed methodology made it convenient to evaluate the sentiments and emotions of the feedback provided by the employees. It states the sentiments of the review and talks about the polarity of that sentiment [12]. The polarity is either positive, negative, or neutral.

2. Proposed Methodology for Employee Feedback Analysis

Sentimental analysis assists to monitor the emotion hidden behind any sentence. For the Sentimental Analysis of the Employee Reviews, the author proposed an Employee Feedback Model which made it easier to analyze the sentiments or emotions of the reviews and the polarities of the sentiments [15-16]. To test this feedback model TCS (Tata Consultancy Services) Employee Review data was taken into consideration which consisted of Indian as well as International Reviews. The steps followed in the development of the model are explained below:
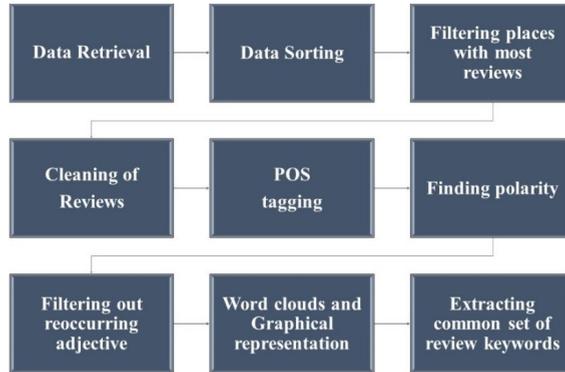
Fig. 1: Steps for proposed Methodology

Step 1: Data Retrieval: The data was taken from Kaggle on the 25th of May 2022. The size of the data is 5 Mib.

Step 2: Sorting of Data: The data were arranged in meaningful order to get the clarity of the highest and lowest number of reviews.

Step 3: After the completion of sorting data, the places with the greatest number of reviews were taken into consideration for the further process.

Step 4: Next step was to clean the data as the elimination of the null, void and invalid data was important to move ahead in the evaluation.

Step 5: POS Tagging: The separation and tagging of the Adjective, Noun, and Verb present in the reviews was done in this particular stage.

Step 6: After the tagging is done the polarity of reviews was calculated precisely through Ruled Based Analysis.

Step 7: In this step the most occurring adjectives were filtered out from the cleaning reviews to make Word clouds.

Step 8: At last graphs and tables were made for the deep and thorough analysis of the polarities of the reviews**.**

2.1 Dataset

The following data was taken from Kaggle on 25th May 2022. The size of the Dataset is 5 MB. The analysis was done after some cleaning and sorting of data from the original dataset.

Janhvi Juyal[1] , Sonakshi Vij[2] , Deepali Virmani[3]

Table 1. Dataset Elucidation

Dataset Title	Description	Source	Size	Date of Retrieval	URL
TCS employee Reviews	The dataset consists of TCS employee reviews from all across the world.	Kaggle	5 MB	25th May 2022	https://www.kaggle.com/datasets/ashwini1619/tcsglassdoorreviews

 Table 1 represents the description of the dataset used in the paper. This Dataset consists of TCS employee reviews from India as well as reviews from all across the globe.

The mentioned dataset was initially of 46132 rows which consisted of lots of null and void values. To make the analysis much more precise the data was filtered and reduced first. The cleaned and sorted data consisted of places with the greatest number of reviews. The sentimental analysis was done on the column named "Review Headline" in the dataset. The column was composed of all the reviews that were written by employees across the world.

2.2 Results & Implementation

The model was proposed for the Sentimental Analysis of employee reviews. The dataset consisted of positive as well as negative feedback which was provided by the employees regarding TCS. The model will help to predict the sentiments of the reviews precisely and the polarities of those sentiments. This technique of identifying the sentiments of the employee reviews helped to judge the employee feedback so that the company would understand the state of mind of every employee. This positively contributes to work performance and engaging behavior hence promoting the growth of the company [9-11]. The table given below provides the data of the total number of reviews (positive or negative) given by the employees nationally and internationally.

Table 2. Total Number of employee reviews in India vs International

Place	Total Number of reviews
India	18927
International	841

Table 2 depicts the total number of employee reviews nationally and internationally. Clearly as illustrated in the above-shown table the number of national reviews is much more than then the number of international reviews which implies most of the reviews are given by employees living in India.

Table 3. Total number of reviews from each of the top places in India

Place	Total Number of reviews
Bangalore	4888
Calcutta	1553
Chennai	3995
Gurgaon	1256
Mumbai	4365
Pune	2708

Table 3 illustrates the number of employee reviews per place nationally. From the data provided by the table, it is understandable that Bangalore has the highest number of reviews which is 4888 which implies it has more placements than any other place.

Table 4. Total number of reviews from each of the top places internationally

Place	Total Number of Reviews
Budapest	160
Guadalajara	116
London	177
New York	189
Toronto	193

Table 4 shows the number of employee reviews per place internationally. From the data provided in the above table, it is clear that London has the highest number of employee reviews which is 193 which implies that it has more TCS employees than any other place. With the help of the proposed model, the polarity of reviews is determined precisely which is represented by the graphs mentioned below

2.3 The polarity of Indian Reviews:
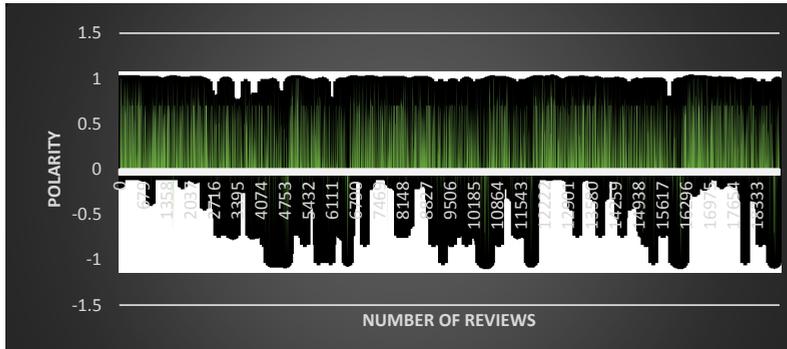
Fig. 2: Graph manifesting polarity of all the Indian Reviews.

Figure 2 highlights the polarity of reviews calculated for the national locations i.e., from employees of TCS offices in India. The graph depicts that only a few numbers of employees have given negative reviews and the rest are satisfied with the company's conduct towards them.

Janhvi Juyal[1], Sonakshi Vij[2], Deepali Virmani[3]

2.4 The polarity of International Reviews:

Fig. 3: Graph manifesting polarity of all the International Review**.**

Figure 3 illustrates the polarity of reviews calculated for the international locations i.e., from employees of TCS offices all across the world**.** The graph illustrates that even people working outside of the nation are happy with the current situation of the organization.

2.5 World Cloud: India vs International

India:

Fig. 4: Depicts the Word Cloud generated from Indian reviews**.**

The Word Cloud in figure 3 majorly contains reoccurring words in the feedback or reviews like 'good company', 'work-life balance, 'good experience', 'good place', 'job security, etc. which implies that a lot of the reviewers gave favorable feedback. Figure 3 depicts the collection of the Indian reviews and the word cloud made up of the thorough analysis of feedback from each place is shown below:

Bangalore:

Fig. 5: Represents the Word Cloud generated from Bangalore reviews.

Figure 4 represents the control words & keywords that were obtained as the top review comments from Bangalore. The top 5 review keywords are represented by the set S(B) as follows:
S(B)= {"good", "salary", "life balance", "experience", "good company"}

Calcutta:

Fig. 6: Represents the Word Cloud generated from Calcutta reviews.

Figure 6 shows the control words & keywords that were obtained as the top review comments from Calcutta. The top 5 review keywords are represented by the set S(C) as follows:
S(C)= {"good", "experience", "company work", "good company", "great"}

Chennai:

Fig. 7: Represents the Word Cloud generated from Chennai reviews

Figure 7 depicts the control words & keywords that were obtained as the top review comments from Chennai. The top 5 review keywords are represented by the set S(C1) as follows:
S(C1) = {"good company", "good", "experience", "company work", "good place"}

Gurgaon:

Fig. 8: Represents the Word Cloud generated from Gurgaon reviews

Janhvi Juyal[1], Sonakshi Vij[2], Deepali Virmani[3]

Figure 8 manifests the control words & keywords that were obtained as the top review comments from Gurgaon. The top 5 review keywords are represented by the set S(G) as follows:
S(G) = {"good", "good company", "experience", "great company", "life balance"}

Mumbai:

Fig. 9: Represents the Word Cloud generated from Mumbai reviews

Figure 9 represents the control words & keywords that we obtained as the top review comments from Mumbai. The top 5 review keywords are represented by the set S(M) as follows:
S(M) = {"good company", "company work", "experience", "good", "great"}

Pune:

Fig. 10: Represents the Word Cloud generated from Pune reviews

Figure 10 represents the control words & keywords that were obtained as the top review comments from Pune. The top 5 review keywords are represented by the set S(P) as follows:
S(P) = {"good company", "company work", "experience work", "good", "work-life"}

International:

Fig. 11: Depicts the Word Cloud generated from international reviews.

'Good company', 'experience', 'opportunity', and 'great company', are some words repetitively shown in Figure 11, which depicts those employees who are satisfied with the current situation of the organization globally. The above-displayed word cloud is the combination of feedback from different places internationally. To study in more detail the word clouds of different places are mentioned below.

Budapest:

Fig. 12: Represents the Word Cloud generated from Budapest reviews

Figure 12 represents the control words & keywords that were obtained as the top review comments from Budapest. The top 5 review keywords are represented by the set S(B) as follows:
S(B) = {"good", "great", "experience", "environment", "grow"}

Guadalajara:

Fig. 13: Represents the Word Cloud generated from Guadalajara reviews

Figure 13 represents the control words & keywords that were obtained as the top review comments from Guadalajara. The top 5 review keywords are represented by the set S(G) as follows:
S(G) = {"good", "good company", "experience", "great", "nice"}

London:

Fig. 14: Represents the Word Cloud generated from London reviews

Janhvi Juyal[1] , Sonakshi Vij[2] , Deepali Virmani[3]

Figure 14 represents the control words & keywords that were obtained as the top review comments from London. The top 5 review keywords are represented by the set S(L) as follows:
S(L) = {"good", "good company", "nice", "great", "learning environment"}

New York:

Fig. 15: Represents the Word Cloud generated from New York reviews

Figure 15 represents the control words & keywords that were obtained as the top review comments from New York. The top 5 review keywords are represented by the set S(NY) as follows:
S(NY)= {"good", "great", "experience", "opportunity", "good company"}

Toronto:

Fig. 16: Represents the Word Cloud generated from Toronto reviews

Figure 16 represents the control words & keywords that were obtained as the top review comments from New York. The top 5 review keywords are represented by the set S(T) as follows:
S(T)= {"good work", "great", "experience", "best", "learning"}

3. Results
The result of the above-mentioned Employee Feedback Model gave clarity about the polarity of sentiments of the feedback and the reviews provided by the employees concerning the organization.
The below-displayed graph depicts the average polarities of different places in India.

Fig. 17: Representation of Average Polarities of different places in India

Figure 17 depicts the average polarities of the top places in India. Pune is the place with the highest positive polarity of .335461 and Calcutta is the place with the lowest polarity of 0.27188.
The below-displayed graph depicts the average polarities of different international places.
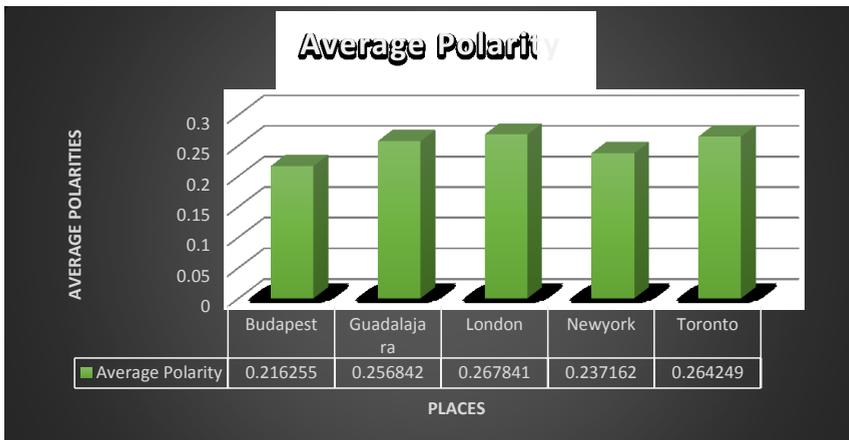
Fig. 18: Representation of Average Polarities of individual international places

Figure 18 illustrates the average polarities of top places internationally. Toronto is the place with the highest positive polarity of .264249 and Budapest is the place with the lowest polarity of 0.216255.

*-The below-displayed graph depicts the comparison among the average polarity of India and International.

Janhvi Juyal[1] *, Sonakshi Vij*[2] *, Deepali Virmani*[3]
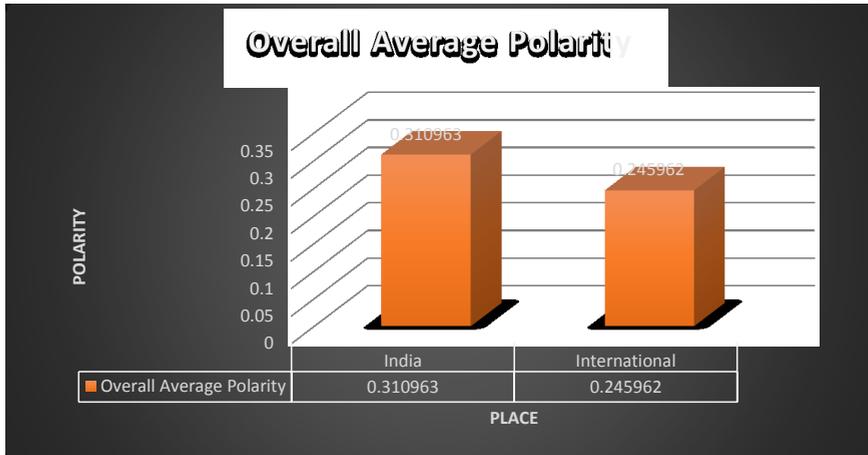
Fig. 19: Representation of India vs International average polarity

Figure 19 represents the overall average Indian polarity as well as the average international polarity of the dataset. It shows the comparison between the overall average Indian as well as international polarities. Regarding the above graph, it is depicted that the Average Indian polarity is 0.310963 and the average international polarity is 0.245962. This polarity illustrates that people who are working with TCS are satisfied and happy with the present situation and condition. The overall sentiments of all the reviews are represented by the set S(R) as follows:

S(R)= {"Good Company**"**, "Experience**"}**

"Good company" and "experience" are the set of words or keywords that occurred repeatedly in any place's review, be it national or international. Being the most occurring words, they were seen in almost every word cloud with a very high frequency which implies that employees stated TCS as a good company and a company that provides individuals with professional experience. This result identifies which aspect of the company is more preferred by the current employees which help the company to work and develop a better environment or working space for the employees serving the company. This would automatically and spontaneously promote the growth and professional development of the company. This signifies that TCS being such a successful multinational IT company not only provides people with consultancy services but at the back end takes a lot of care of its employees and thinks about their benefit.

4 Conclusion

The feedbacks retrieved from employees are crucial for any organization as it assists in the long term to increase the productivity of its employees and make them feel significant. While positive reviews depict the strong key points of any company, negative feedbacks provide constructive criticism. The authors have proposed a novel method based on sentiment analysis to generate an analysis method for an employee feedback system. The proposed method was tested on the TCS employee review data which consisted of Indian as well as International Reviews. The common set of review keywords extracted using the proposed approach was "good company" and "experience". The results show a positive inclination towards the company and represent that people are generally happy and satisfied with the organization. In the future, this method could be applied to other company datasets as well.

References

1. Kumar, A., & Jain, R. (2015, October). Sentiment analysis and feedback evaluation. In *2015 IEEE 3rd International Conference on MOOCs, Innovation, and Technology in Education (MITE)* (pp. 433-436). IEEE.

2. Altrabsheh, N., Cocea, M., & Fallahkhair, S. (2014, November). Sentiment analysis: towards a tool for analyzing real-time student feedback. In *2014 IEEE 26th international conference on tools with artificial intelligence* (pp. 419-423). IEEE.

3. Park, E., Kang, J., Choi, D., & Han, J. (2020). Understanding customers' hotel revisiting behavior: sentiment analysis of online feedback reviews. *Current Issues in Tourism*, *23*(5), 605-611.

4. Kastrati, Z., Dalipi, F., Imran, A. S., Pireva Nuci, K., & Wani, M. A. (2021). Sentiment analysis of students' feedback with NLP and deep learning: A systematic mapping study. *Applied Sciences*, *11*(9), 3986.

5. Van Nguyen, K., Nguyen, V. D., Nguyen, P. X., Truong, T. T., & Nguyen, N. L. T. (2018, November). UIT-VSFC: Vietnamese students' feedback corpus for sentiment analysis. In *2018 10th international conference on knowledge and systems engineering (KSE)* (pp. 19-24). IEEE.

6. Nasim, Z., Rajput, Q., & Haider, S. (2017, July). Sentiment analysis of student feedback using machine learning and lexicon-based approaches. In *2017 international conference on research and innovation in information systems (ICRIIS)* (pp. 1-6). IEEE.

7. Moghaddam, S. (2015, March). Beyond sentiment analysis: mining defects and improvements from customer feedback. In *European conference on information retrieval* (pp. 400-410). Springer, Cham.

8. Lundqvist, K., Liyanagunawardena, T., & Starkey, L. (2020). Evaluation of student feedback within a MOOC using sentiment analysis and target groups. *International Review of Research in Open and Distributed Learning*, *21*(3), 140-156.

9. Chauhan, G. S., Agrawal, P., & Meena, Y. K. (2019). Aspect-based sentiment analysis of students' feedback to improve the teaching-learning process. In *Information and communication technology for intelligent systems* (pp. 259-266). Springer, Singapore.

10. Shakhovska, K., Shakhovska, N., & Veselý, P. (2020). The sentiment analysis model of services Providers' feedback. *Electronics*, *9*(11), 1922.

11. Ullah, M. A. (2016, December). Sentiment analysis of students feedbacks: A study towards optimal tools. In *the 2016 International Workshop on Computational Intelligence (IWCI)* (pp. 175-180). IEEE.

12. Singh, L. K., & Devi, R. R. (2021). Student feedback sentiment analysis: A review. *Materials Today: Proceedings*.

13. Ramanathan, V., & Meyyappan, T. (2019, January). Twitter text mining for sentiment analysis on people's feedback about oman tourism. In *2019 4th MEC International Conference on Big Data and Smart City (ICBDSC)* (pp. 1-5). IEEE.

14. Asghar, M. Z., Ullah, I., Shamshirband, S., Kundi, F. M., & Habib, A. (2019). Fuzzy-based sentiment analysis system for analyzing student feedback and satisfaction.

15. Pandey, P., & Soni, N. (2019, February). Sentiment analysis on customer feedback data: Amazon product reviews. In *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)* (pp. 320-322). IEEE.

16. Kumar, V. V., Raghunath, K. M., Muthukumaran, V., Joseph, R. B., Beschi, I. S., & Uday, A. K. (2022). Aspect-based sentiment analysis and smart classification in uncertain feedback pool. *International Journal of System Assurance Engineering and Management*, *13*(1), 252-262.

Janhvi Juyal[1] , Sonakshi Vij[2] , Deepali Virmani[3]

17. Chakravarthi, Bharathi Raja, et al. "Dravidiancodemix: Sentiment analysis and offensive language identification dataset for Dravidian languages in code-mixed text." *Language Resources and Evaluation* (2022): 1-42.

18. Li, H., Chen, Q., Zhong, Z., Gong, R. and Han, G., 2022. E-word of mouth sentiment analysis for user behavior studies. *Information Processing & Management*, *59*(1), p.102784.