



# Analysis of Human Resources Attrition: A Thematic and Sentiment Analysis Approach

Punamkumar Hinge<sup>(✉)</sup> , Abhijeet Thakur, and Harshal Salunkhe 

School of Business and Management, Christ University, Bangalore, India  
punamkumar.hinge@gmail.com

**Abstract.** In this paper, factors are analyzed for employee attrition to find the main reasons why employees choose to resign and suggested how Artificial Intelligence can be developed to understand and predict future attrition reasons by using thematic and sentiment analysis. The prime objective of study is why people leave the organization and how it can be predicated in advance by using NLP (Natural language processing) AI. Different weights are given for different factors for attrition reasons based on the study conducted for one of the auto-component manufacturing company in Chakan MIDC, Pune, India, high weight factors will be alarming signals for the organizations to be proactive to correct or rectify so that they can avoid attrition. Study can help which valuable employees will leave organization for different reasons in advance. The investigation was done to determine what element had the biggest impact on employee attrition. With the help of the paper, new work policies can be created that benefit both the business organisation and the employee. It might be viewed as a reflection of the employees' working environments.

**Keywords:** artificial intelligence · employee attrition; prediction model · employee sentiment analysis · thematic analysis · natural language processing

## 1 Introduction

Attrition is defined as voluntary or involuntary resignation of the employees also Attrition is defined as an employee resigning or retiring from a company [1]. These days, all firms view employee turnover (also known as attrition) as their top concern due to its negative consequences on workplace productivity and meeting deadlines for corporate goals. An organization should make it a duty to reduce employee attrition in order to consistently have a higher competitive advantage over its competitors. By analyzing the employee data on various aspects like career plan, promotion, recruitment, age, performance appraisal data, training data organization can identify different reasons of attrition then organizations can develop algorithm by using machine learning so that it can suggest possible attrition of employees in future and its reason based on different factors so that companies can take predictive measures to improve human resources policies [2].

## 2 Literature Review

The use of artificial intelligence within organizations affects a company's decision-making processes in a number of ways. Human resources (HR) have received more attention recently because the caliber and aptitude of employees are a growth element and a true competitive advantage for businesses [3]. In fact, artificial intelligence is now beginning to influence business decisions regarding their employees after becoming more widely used in the sales and marketing sectors, with the intention of basing HR management decisions on the analysis of objective data rather than subjective considerations. The employee's specialization and continuity of work become crucial for businesses where workers undertake more activities that are specialized. By applying predictive models, which use employee data gathered by the company over the previous years, the application of artificial intelligence in the field of HR enables businesses to turn data into knowledge, eliminating important difficulties and optimizing all HR tasks. Companies devote a lot of effort and money to hiring new employees and providing them with the necessary training in order to meet their strategic objectives. Employees, therefore, represent a meaningful investment for organizations (to a greater or lesser amount). When an employee quits the company, it forfeits not only the services of that valuable individual but also the time, money, and resources expended in finding, selecting, and preparing them for their unique jobs. Attrition is defined as an employee resigning or retiring from a company [1]. There are various reasons of employee attrition e.g. below standard of living payment, Bad treatment of manager, long working hours, no promotional opportunities, religious groupism, regional groupism, non supporting seniors, pay discrimination for same rank, appraisal biasness, irregular salary, job insecurity, Away from city, excessive work load, Unhygienic conditions, unregularly increments, no training to upgrade skills, no career plan, not safe for woman, Slow growth in career, manipulated feedback, poor grievance mechanism, poor employee welfare, no compensatory off for worked on holidays or Sundays, no proper leaves, denying from taking leaves, improper system, poor work culture, harassments at workplace, Gender biasness in promotion, fraudsters practices of superiors, lack in futuristic plan of organization, lack of company vision, faulty appraisal system. Costs associated with the employee attrition cycle in human resources, knowledge loss, low employee morale, and organizational culture [4]. Employee attrition can develop into a severe problem because it can harm an organization's ability to compete. The foundation of any business is its satisfied, driven, and loyal employees, who in turn affect an organization's productivity [5].

The use of classification algorithms can assist HR management by enabling the implementation of staff management support technologies in the business. People management skills have strong relationship with respect to attrition. Managers bad treatment to employees leads to higher attrition rate or dissatisfaction [6]. Finding a well-trained and experienced staff is a difficult endeavor for any firm, but replacing such personnel is even harder. If the individual hired to replace the long-serving employee is unable to maintain a positive relationship with the client, it may have an adverse effect on the company's revenue when the long-serving staff leave [7]. Employee attrition (turnover) is expensive for any business, and it could eventually reduce productivity [8]. Companies develop new employee retention programs every year in an effort to prevent staff attrition, which has a detrimental impact on the growth of businesses, these strategies include

stock options that mature after a specific amount of time, improved training, retention bonuses, and better training. Therefore, identifying the key cause of attrition will enable organizations in developing more focused retention strategies [9]. Employee attrition is a major problem for firms, particularly when trained, technical, and key employees leave in search of better opportunities elsewhere. Analyzing the prevalent causes of employee attrition using data on current and former employees can help to predict attrition where organizational can use artificial intelligence and machine learning [10]. Artificial intelligence (AI) developments have caused the corporate environment to change quickly. AI integration in human resources will make it easier to analyze, forecast, and diagnose the problems that firms encounter and will aid in improving employee-related decisions [11]. Human language is complex by nature. A technology needs to comprehend grammatical conventions, meaning, context, as well as colloquialisms, slang, and acronyms used in a language in order to comprehend human speech. Computers are supported by natural language processing (NLP) algorithms, which mimic human comprehension of linguistic information, including unstructured text data [12]. The use of language affects almost every area of business. We produce an incredible amount of unstructured data as a result (e.g., PDFs, business documents, emails, videos, etc.). 80% or 90% of corporate data is actually unstructured. This information is priceless and can offer crucial business insight. However, NLP is required to unlock it [13].

### 3 Research Methodology

#### 3.1 Algorithm Used

**Natural Language Processing (NLP) algorithms** such as Keyword Extraction, Topic Modeling, Word Cloud, Sentiment Analysis, and Tokenization were used.

Following stages then used for conducting the research:

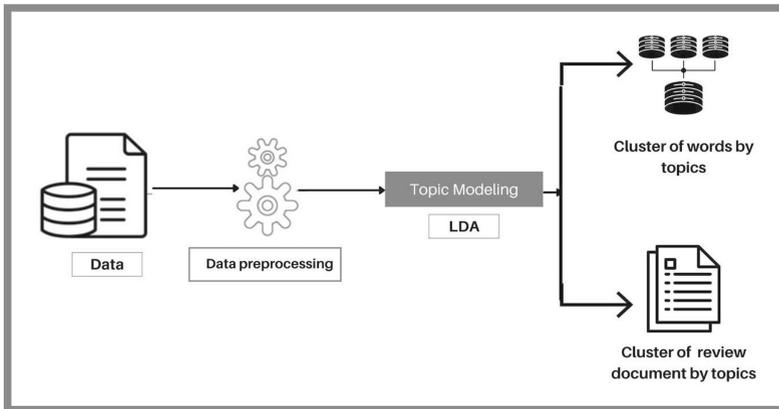
*Stage 1: Data Collection:* Since employee attrition is internal corporate data that is challenging to collect and some of it has a certain level of confidentiality, my research used the data set of 109 employees made available by **one auto-component manufacturing organization located at Pune MIDC, India.**

*Stage 2: Data pre-processing:* Data collected in the first stage was imported to Orange application software where data 1. Transformed into lower cases and removed tags and URLs. 2. Tokenization: Under Tokenization, data was splitted in to regular expressions and would be using or keeping only words. 3. Normalization: Wordnet Lemmatization technique was used to process the data set and extracted some contextual meaning and text out of the document. 4. Filtering: In this process, stop words are removed in the English language. This step processes the selection of words and removed unwanted words, symbols, etc. Stop word is used to remove (and, or...in) from the text document.

*Stage 3: Topic Modeling:* Topic Modeling with Latent Dirichlet Allocation Process (Fig. 1) is conducted for further analysis.

*Stage 4: Thematic Analysis:* Thematic analysis is used to understand the meaning of data so as data can be categorized into topics and divide into themes, which helps to analyzed, and interpret data.

*Stage 5: Sentimental Analysis:* Using Liu Hu method, Sentiment Analysis predicts sentiment for each document in a corpus. Here we used Liu Hu model.



**Fig. 1.** Topic modeling Process

## 4 Result Analysis and Interpretation

### 4.1 Topic Modeling

109 employees reviews from one auto-component manufacturing organization from MIDC, Chakan, Pune, India were taken where the extracted reviews were pre-processed by transforming the text to lower case, tokenizing the data, and filtering the data by removing stop words such as ‘the’, ‘an’, ‘a’, numbers from the text and include only meaningful words. Topic modeling was carried out on these reviews. Latent Dirichlet Allocation method of topic modelling was used to define topics. The wordcloud was formed representing keywords. Keywords were also extracted separately. The sentiment analysis on this topic was carried out using VADER method, which further analyses the statements as positive, negative and neutral.

#### Wordcloud

In order to identify the keywords that made sense and were important from the perspective of the company to understand which aspects of their employment need to be improved further to enhance the employee experience, the extracted reviews were pre-processed and topic modeling was applied. A word cloud represents words in the form of a cloud, arranged based on a weight assigned to each word Table 1. Accords weights to words according to their number of occurrences, arranged in descending order of their weight. In the center of the cloud, the words with the highest weight are in bold and larger font size, and as the weight of the words decreases, the font size decreases (Table 2).

**Table 1.** Word Frequency Table/Weights of the Word

<b>Weight</b>	<b>Word</b>	<b>Weight</b>	<b>Word</b>
18	Poor	3	leaves
12	Career	3	better
7	Work	3	life
6	Working	3	work life
6	lack	3	managers
6	management	3	retirement
6	benefits	3	career growth
5	Job	3	development
4	employees	3	training development
4	opportunities	2	change
4	training	2	long
3	growth	2	hours
3	bad		
3	pay		
3	Conditions		

**Table 2.** Sentiment Analysis

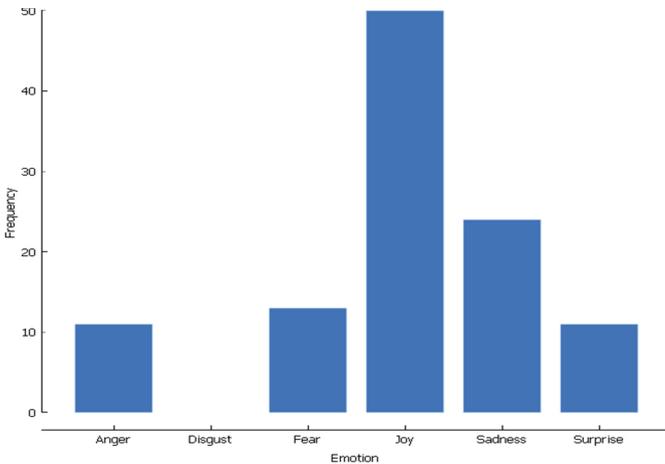
<b>Sentiments</b>	<b>No. of employees</b>	<b>Percentage</b>
Anger	11	10%
Disgust	0	0%
Fear	13	12%
Joy	49	45%
Sad	25	23%
Surprise	11	10%

From the above word frequency table it is found that there are some words which has highest weightages which are the major reasons resulting into the attrition e.g. poor, career, management, working etc. it simply says that poor work culture or poor management treatment to the employees or not having career opportunities in the existing organization, lack of proper benefits, poor pay, lack of training and development, poor leave policies, poor work life balance is resulting into the higher attrition rates. Company should focus on these higher weightage words, which are the major reasons for the attrition. Company can predict the attrition well in advance and correct it for future to avoid attrition by revising the human resources strategies (Fig. 2).



<b>Training &amp; Development</b>		<b>Career Growth</b>	<b>Future Prospects</b>
Career Poor Growth Training Development		Job Change Job Security Promotion Career change	Better opportunities lack of job groupism opportunities
<b>Employee Welfare &amp; Relation</b>			
Managers Poor Leave Policy Manager employee relation		Employee benefits Talent Working hours	

**Fig. 3.** Thematic analysis



**Fig. 4.** Sentiment Analysis

### 4.3 Sentiment Analysis

In sentiment analysis, process bar chart is used to understand the general reviews of the 109 employees regarding company, which then became the basis for their attrition reasons. In the analysis, it is found that 10% employees have shown anger towards current employee condition, means they are not satisfied in the company. 0% employees shown disgust atmosphere or unpleasant experience in the current company. 12% employee found fear to continue this organization. 45% employees shown joy working in this company. 23% are sad in their employment in this organization. 10% employees are surprised mean they were not seen such type of working environment in their former organizations (Fig. 4).

## 5 Conclusion

I want to make some suggestions to organization that company should train well their managers to behave well with their subordinates to have respectful atmosphere. Organizations should improve job satisfaction by increasing work culture, reducing overburden of work, avoiding excessive working hours, setting proper measurable standards for performance appraisal, paying at par standard of living salary, providing career opportunities to employees, allowing employees in management decision making, providing proper & sufficient paid leave policies, allowing them to take proper breaks for their personal reasons so that employees will become productive and motivated to stay in the organizations in a long term. Company can develop AI algorithm on the basis of the above thematic and sensitivity analysis which can predict attrition and prescribe organization necessary measures to be taken to avoid the future attrition.

## References

1. S. S. Alduayj and K. Rajpoot, "Predicting Employee Attrition using Machine Learning," Proc. 2018 13th Int. Conf. Innov. Inf. Technol. IIT 2018, pp. 93–98, Jan. 2019, doi: <https://doi.org/10.1109/INNOVATIONS.2018.8605976>.
2. S. Yang, M. I. preprint [arXiv:2012.01286](https://arxiv.org/abs/2012.01286), and undefined 2020, "IBM Employee Attrition Analysis," arxiv.org, Accessed: Oct. 29, 2022. [Online]. Available: <https://arxiv.org/abs/2012.01286>
3. F. Fallucchi, M. Coladangelo, R. Giuliano, and E. W. De Luca, "Predicting Employee Attrition Using Machine Learning Techniques," Comput. 2020, Vol. 9, Page 86, vol. 9, no. 4, p. 86, Nov. 2020, <https://doi.org/10.3390/COMPUTERS9040086>.
4. I. Setiawan, S. Suprihanto, ... A. N.-I. C., and undefined 2020, "HR analytics: Employee attrition analysis using logistic regression," iopscience.iop.org, <https://doi.org/10.1088/1757-899X/830/3/032001>.
5. D. Kamath, ... D. J.-I. J. T. S. R. D., and undefined 2019, "Machine Learning Approach for Employee Attrition Analysis," academia.edu, Accessed: Oct. 29, 2022. [Online]. Available: [https://www.academia.edu/download/59804527/17\\_Machine\\_Learning\\_Approach\\_for\\_Employee\\_Attrition\\_Analysis20190620-80418-1hdg8n.pdf](https://www.academia.edu/download/59804527/17_Machine_Learning_Approach_for_Employee_Attrition_Analysis20190620-80418-1hdg8n.pdf)
6. M. Hoffman and S. Tadelis, "People management skills, employee attrition, and manager rewards: An empirical analysis," J. Polit. Econ., vol. 129, no. 1, pp. 243–285, Jan. 2021, <https://doi.org/10.1086/711409>
7. K. Sehgal, H. Bindra, A. Batra, R. J.-I. in computer science and, and undefined 2019, "Prediction of employee attrition using GWO and PSO optimised models of C5. 0 used with association rules and analysis of optimisers," Springer, Accessed: Oct. 29, 2022. [Online]. Available: [https://link.springer.com/chapter/https://doi.org/10.1007/978-981-13-7082-3\\_1](https://link.springer.com/chapter/https://doi.org/10.1007/978-981-13-7082-3_1)
8. S. Yadav, A. Jain, D. S.-2018 I. 8th I. Advance, and undefined 2018, "Early prediction of employee attrition using data mining techniques," ieeexplore.ieee.org, 2018, Accessed: Oct. 29, 2022. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/8692137/>
9. D. K. Srivastava and P. K. Tiwari, "An analysis report to reduce the employee attrition within organizations," J. Discret. Math. Sci. Cryptogr., vol. 23, no. 2, pp. 337–348, Feb. 2020, <https://doi.org/10.1080/09720529.2020.1721874>
10. R. Shankar, ... J. R.-2018 ieeec, and undefined 2018, "Prediction of employee attrition using datamining," ieeexplore.ieee.org, 2018, <https://doi.org/10.1109/ICSCAN.2018.8541242>.

11. M. Arora, A. Prakash, ... A. M.-... C. on D., and undefined 2021, "HR Analytics and Artificial Intelligence-Transforming Human Resource Management," ieeexplore.ieee.org, Accessed: Nov. 02, 2022. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/9682325/>
12. X. Qiu et al., "Pre-trained models for natural language processing: A survey," Springer, Accessed: Nov. 18, 2022. [Online]. Available: <https://link.springer.com/article/https://doi.org/10.1007/s11431-020-1647-3>
13. D. Khurana, A. Koli, K. Khatter, and S. Singh, "Natural language processing: state of the art, current trends and challenges," *Multimed. Tools Appl.*, 2022, doi: <https://doi.org/10.1007/S11042-022-13428-4>.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

