



# A Study on Employee Attrition using Artificial Neural Networks

B.Sindhura<sup>1\*</sup> | S.V.N.L.Lalitha<sup>2</sup> | Mohammed Sumeiruddin<sup>3</sup>

<sup>1</sup>Assistant Senior Project Engineer, Vijay Nirman Company Pvt. Ltd., Visakhapatnam, India.

<sup>2</sup>Professor, Department of EEE, KLEF Deemed to be University, Guntur, India.

<sup>3</sup>Department of Civil Engineering, JNTUH College of Engineering, Hyderabad, Nalgonda, India.

\*Corresponding Author Email Id: [sindhubs@gmail.com](mailto:sindhubs@gmail.com)

## To Cite this Article

B.Sindhura, S.V.N.L.Lalitha and Mohammed Sumeiruddin. A Study on Employee Attrition using Artificial Neural Networks. International Journal for Modern Trends in Science and Technology 2022, 8(07), pp. 158-163. <https://doi.org/10.46501/IJMTST0807023>

## Article Info

Received: 04 June 2022; Accepted: 30 June 2022; Published: 06 July 2022.

## ABSTRACT

*This paper aims at studying the variables leading to employee attrition by characterizing the data related to an individual employee of any Organization. Application of data analytics for assessing the human resource (HR) management is emerging out as a key area. Owing to the interrelationship of several variables leading to employee attrition, research is now directed towards the machine learning algorithms in HR analytics. An artificial neural network (ANN) has been proved to be a promising tool for recognizing the correlation between various parameters as well as for its learning and adoption capabilities. An attempt has been made to study the employee attrition using the ANN in R, a programming language for statistical Computing and Graphics. Identification of important factors driving the employee attrition in an organization has been done. A case study has been performed considering the variables which would eventually lead to employee attrition in different levels of employees in any organization. The HR data published by the IBM Watson Company has been used for training the neural network. Primary data has been collected through an online survey using snowball sampling technique from employees working in different organization at different levels such as junior/senior executive, team lead, assistant manager/manager, senior manager and above using questionnaire. Based on the thorough analysis of outcomes of the case study, suitable remedial measures to reduce the employee attrition have been presented.*

**KEYWORDS** – Attrition, Artificial Neural Networks, R Programming, HR Analytics

## 1. INTRODUCTION

Attrition, also known as employee turnover, or employee defection, is an industrial term used to describe loss of employees or Man power. Generally, attrition means how many employees left the organization other than the members who have retired or have been terminated by the organization. Employee attrition is one of the major concerns in any organization.

There are several visible and invisible costs associated with employee attrition. The problem of attrition has gained significance in the wake of lack of skilled employees to meet the ever-changing technology development.

Employees leave organizations for many different reasons. Sometimes it is the attraction of a new job or the prospect of a rewarding period outside the workforce

which 'pulls them. On other occasions they are 'pushed' (due to dissatisfaction in their present jobs) to seek alternative employment [1]. Several factors which lead to lead to employee attrition may be classified into Insufficient Incentives, Low Job Satisfaction, Poor Job Performance, Labour Market Conditions and Other Role Commitments [2]. Several factors which affect employee attrition have been identified by a survey conducted by Rashmi Priya (Employee Attrition: 2008: Google) [3].

It is useful to calculate a separate figure for voluntary attrition and to consider some of the more complex employee attrition indices, which take account of characteristics such as seniority and experience (Scuff, 2010) [4].

Attrition represents significant costs to technology and business process outsourcing (BPO) companies. High attrition rates drive up training costs, and increase human resources recruiting and productivity costs. They also increase the prospect of customer service complaints or quality problems and create substantial continuity problems for longer-lived projects. It is further pointed out that a company has a training period of 3 to 6 months. During this time an employee is not fruitful for the company. If an employee leaves the company after training, the company suffers a big loss in terms of money as well as trained workforce [5]-[7].

Machine learning, a branch of artificial intelligence, was originally employed to develop techniques to enable computers to learn. Machine learning algorithms have become an indispensable tool for any analysis. Artificial Neural Networks (ANN) have emerged out as a promising tool for pattern recognition and as excellent classifiers. Several studies have been done in various fields on the application of ANN. Application of text mining, data mining and statistics using SPSS software has been reported in the literature [8]-[9].

The role of HRM has changed over time from the administrator of legal and mandatory human resource practices to more of a supporter of value creation and business strategy. Since the early 20th century, the collection and processing of employee data has been the central element in HRM. The HR related data is still often defined as hard to define, difficult to measure and not critical when it comes to business strategy. In addition to the data-driven measuring and decision-making, it has been argued that HR needs to embrace an outside-in approach. This means that if HR

wants to fully validate their role as a true strategic partner, they do not only have to be able to make the data-driven decisions, but also need to relate these decisions to what happens outside the organisation, on the business context and external stakeholders, particularly the customers [10]-[11].

A recent Deloitte report stated that a third of the company's state that HR analytics is under active development, but only eight per cent of them claim to have strong analytics capabilities (Deloitte, 2015) [12]. In addition, they show little sign of improvement compared to the previous year, indicating thereby that the expectation of HR analytics may not have yet evolved into reality.

HR analytics is becoming the order of the day in view of large data. Predictive analytics are a rapidly upcoming trend in Human Resources (HR). Even though a lot of people talk about predictive analytics in HR, hardly any organizations apply them to their workforce. HR predictive analytics are real game changer for HR departments in near future. Predictive analytics are everywhere. It is in its essence a technology that learns from existing data and it uses this to forecast individual behaviour. This means that predictions are very specific. Predictive analytics involve a set of various statistical (data mining) techniques used to predict uncertain outcomes. Attrition Analytics is an interesting area of research in HR analytics and it gives a very good opportunity for the analysis of the reasons for the improvement of the business by retaining their employees for longer time and make them more productive (Cengiz et.al 2015) [13].

## 2. MATERIAL AND METHOD

Exploratory-Descriptive Research Design is used in this paper. Even though each organization has developed its own HR management strategies to retain their employee for longer time, the problem of attrition is not completely resolved in many organizations and there is a vast scope for further research. Development of machine learning algorithms for employee attrition has been made using Artificial Neural Networks. For designing the neural network model, R - a programming language for statistical Computing and Graphics has been explored.

Snowball Sampling technique has been used in the present study to collect the primary data through

Google docs. The respondents include employees working in different companies. A total of 160 employees have been contacted online and out of them 123 have responded who are working at four different levels such as Junior/Senior Executive (31), Team Lead (30), Assistant Manager /Manager (29) and Senior Manager and Above (33).

Secondary data has been used for training the Artificial Neural Network (ANN) as it needs the past history of employees who have left the organization. The data will be available only if the companies conduct exit interviews or online survey of the employees who have already left. This data has been drawn from IBM Watson Analytics data base being used for HR analytics in data science. This data has been retrieved from the website published by IBM. This data also has been used for testing the SPSS package and the R Studio packages. The secondary data is an exhaustive data with a sample size of 1440 employees with around 36 variables describing the employee characteristics ranging from age, gender, education field, training times, employee satisfaction etc. The data set also contains the data about employee attrition in terms of "Yes" or "No".

For the present study, initially the variables which affect employee attrition most have been identified beforehand. Dependent variable is Employee Attrition. A set of specific independent variables have been taken and then they are given as inputs to the neural network with the attrition as a single output variable.

The list of independent variables is as follows

1. Educational Qualification
2. Designation
3. Marital Status
4. Gender
5. Age
6. Education Field
7. Distance From Home
8. Business Travel
9. Total Working Years
10. Number of Companies Worked earlier
11. Years At This Company
12. Years With Current Manager
13. Years Since Last Promotion
14. Relationship Satisfaction
15. Job Satisfaction
16. Environment Satisfaction
17. Monthly Income

18. Training Frequency

19. No. of years in current role

20. Work Life Balance

### 3. ANN ARCHITECTURE

Several parameters influencing employee attrition have been identified and given as input variables to the ANN. The neural network was trained and then tested using the secondary data and primary data as well. IBM Watson analytics data has been used to train the neural network. In this research ANN architecture has been finalized after several trial and errors. A total of 20 input neurons representing the 20 independent variables have been used in this work. For normalization of the numeric values standard normalization procedure has been used. A simple technique of converting textual data to numerical data has been used in this research work. Too many hidden neurons would hamper the training time and may result in large processing time. Too less a number of hidden neurons may not establish proper relation between input and output variables. The number of hidden neurons has been chosen after several trial and errors in this research work as 5. The dependant variable is employee attrition and in this research work attrition is taken as the output neuron.

#### 1. Training

Here each training pattern contains a total of 21 variables of which 20 variables are the independent input variable and one is the dependant output variable. Since the training requires the data related to the past employees also in any organization, an attempt has been made in this research work to train the neural network using the secondary data initially and testing it with the primary data. A careful justification of the training patterns is highly essential for proper learning by the neural network. Once the network is properly trained, it can adopt to any situation by its self-adaptive nature. A careful examination of the no. of training patters has been carried out in this research. The number of training patters has been chosen as 400 out of 1440 employee database and care has been taken to include the employees working in different organizations at various levels. Initially the secondary data has been segregated into different segments based on the employee level such as junior/senior executive, team lead, assistant manager/manager and senior manager & above. Then

randomly 100 samples have been chosen from each category thus making the number as 400. Supervised learning has been adopted using the back propagation algorithm in this research work.

To test the effectiveness of the developed ANN model using the primary data, a question has been included. The answer of the respondent to the question that "how long are you going to continue in the same organization" reflects the probability of attrition of the employee. It has been inferred that if the response to the above question is "less than a year", it is a clear indication of attrition. Thus, ANN can be tested for assessing employee attrition using the primary data.

A total of 20 independent variables have been given as input to the ANN to determine the likelihood of employee attrition. The data related to these 20 independent variables has been drawn from the responses of the employee to the questionnaire. A total of 123 responses have been obtained from employees working in different organizations at various levels. The entire primary data has been segregated based on the employee level. Primary data comprises of 31 employees at junior/senior executive level, 30 employees at team lead level, 29 employees at assistant manager/manager level and 33 employees at the level of senior manager and above.

## 2. Testing of ANN Model

All the numerical data given as input to the ANN has been normalized using the standard procedure such that all the values remain between 0 and 1. Secondary data drawn from the IBM Watson Analytics company has been used to train the neural network.

## 3. Findings

One question has been included in the questionnaire to do percentage analysis of the important factors leading to employee attrition in an organization. Analysis of the obtained data as shown in Fig.2. clearly indicates "career growth" as one of major drivers for employee attrition.

**3.1 Respondents opinion on factors leading to Leave the job in %:** Following is the inference from Fig.2. 41% of the employees (51 of 123) may leave their present job for better one where there is an opportunity for faster career growth. So, Career growth plays a

dominant role in employee attrition from the obtained data. At each level like junior/ senior executive, Team Lead, Asst Manager/ Manager and Senior Manager & above it can be inferred that Career growth is more important for every employee irrespective of their age, designation and salary which implies that at every level, employees are more inclined towards developing their career and an opportunity to grow at a faster rate.

Another question is aimed at analysing the reasons to reduce employee attrition and retaining the employees in any organization. The analysis illustrated in pie charts Fig.3. demonstrates the visualization of factors which would help in reducing the attrition rate. By careful analysis of the responses, it can be inferred that company profile and working conditions have major effect on the employees.

**3.2 Respondents opinion on factors leading to continue the job in %:** Following is the inference from Fig.3. 27% (33 of 123) and 29% (36 of 123) employees feel that Company profile and Working Condition are important factors to continue in their present job. According to the percentage analysis at every level also, employees are much concerned about the Company Profile & working condition.

The graph in Fig.4. shows the details of the employees at different levels who are likely to leave based on the information obtained from one more question. The three factors career growth, reporting manager's behaviour and salary have been considered and attrition because of these reasons is shown in Fig.4.

These factors have been identified as main reasons because from the percentage analysis of the factors leading to attrition, these three factors have significant proportion among all the factors and hence the percentage of people leaving due to these reasons has been identified and the values for different level of employees working in different companies have been presented below.

The neural network model which has been developed in this research work has been tested with details of 25 employees which contains all the 10 members who have responded that they would leave the Organization. Remaining 15 employees have been chosen randomly. The results obtained with the neural network model are highly encouraging. The model gave the answer 'YES'

for employee attrition for 8 employees out of 10 employees who have given their response as they will be leaving the organization in less than a year. It is important to note that the model has given answer 'NO' for the remaining members.

#### 4. FIGURES

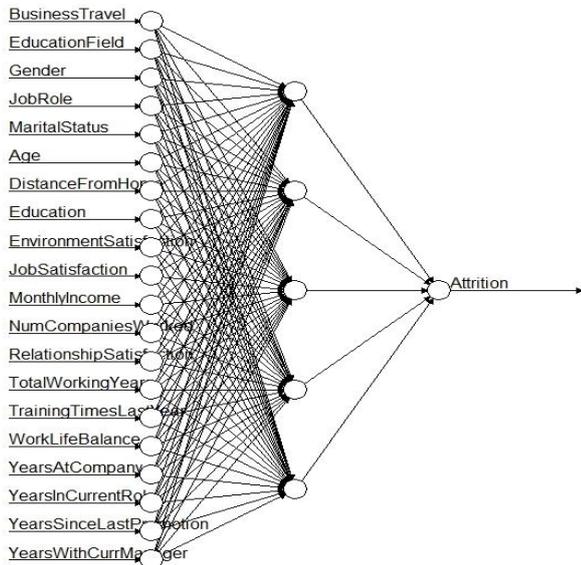


Fig.1. A simplified version of the developed ANN model for assessing employee attrition

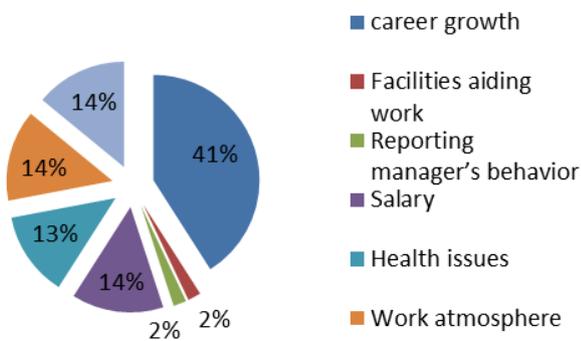


Fig.2. Respondents opinion on factors leading to Leave the job in %

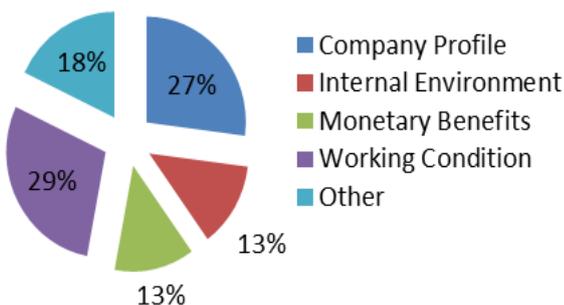


Fig.3. Respondents opinion on factors leading to continue the job in %



Fig.4. Employees working at different levels who are likely to leave

#### 5. DISCUSSION

The developed ANN model gives an "Yes" or "No" answer for the likelihood of an employee to leave an organization based on the input data collected from an employee. However, there are certain personal issues which may lead to employee attrition in which organizations have no role to play. The data related to such employees need to be detected and separated before training the ANN. The data is collected by Snow ball sampling technique and hence cannot be used for hypothesis testing. Researchers may take it up as a future work in this area by taking random sampling.

#### 6. CONCLUSION

In this paper, an attempt has been made to include human cognition through machine learning algorithms to assess one of the most important concerns of HR management i.e., employee attrition. The research reported so far dealt with the analysis of factors affecting employee attrition and their influence on the perception of an employee to leave the organization. In addition to developing a direct method to assess employee attrition using ANN, measures to be taken to reduce employee attrition in an organization have been suggested. HR analytics will be a powerful tool for HR managers to cope up with the technological growth in all walks of life.

#### Conflict of interest statement

Authors declare that they do not have any conflict of interest.

## REFERENCES

- [1] Rankin, N. (2008), The drivers of staff retention and employee engagement. IRS Employment Review. No 901, 1 July. 13pp.
- [2] Randall S. Schuler, Susan E. Jackson, A Quarter-Century Review of Human Resource Management in the U.S.: The Growth in Importance of the International Perspective management revue, vol 16, issue 1, 2005.
- [3] Rashmi Priya 2010, <<http://www.rashmipriya.com/employee-attribution>>
- [4] SUFF, R. (2010) Labour turnover rates and costs: IRS survey 2010. IRS Employment Review. 14 June, 9pp
- [5] Champagne, P., & McAfee, B. (1989). Motivating strategies for performance and productivity: A guide to human resource development. New York: Quorum Books.
- [6] Systemic Coaching for Managers Managing Difficult Employees © Martyn Carruthers.
- [7] INCOMES DATA SERVICES. (2008) Improving Staff Retention. HR Studies, No 863. London: IDS. JANAS, K.M. (2009) Keeping good people during bad times. Workspan. 52,,66-70.
- [8] Yagnanarayana, Artificial Neural Networks, Eastern Economy Edition -1998
- [9] Predicting Employee Attrition Through Data Mining , Abbas Heiat, College of Business, Montana State University, Billings, MT 59102, Proceedings of WDSI 2016 - Las Vegas, NV
- [10] Ulrich, D., & Dulebohn, J. H. (2015). Are we there yet? What's next for HR? Human Resource Management Review, 25(2), 188-204. DOI: 10.1016/j.hrmr.2015.01.004
- [11] Rasmussen, T., & Ulrich, D. (2015). Learning from practice: how HR analytics avoids being a management fad. Organizational Dynamics, 44(3), 236-242.
- [12] Global Human Capital Trends 2015, Leading in the new world of work, Deloitte University Press.
- [13] Intelligent Techniques in Engineering Management: Theory and Applications , edited by Cengiz Kahraman, Sezi Cevik Onar, 2015, Springer