Predicting Employee Attrition using Random Forest with IBM Cloud

Team Pegasuspicious:
Riddhi Gupta, Riya Gupta, Sonali Shripad Shanbhag,
Pendyala Satya Chaitanya

Introduction

Employee attrition became a serious issue regarding a company's competitive advantage. It's very expensive to find, hire and train new talents. It's more cost-effective to keep the employees a company already has. A company needs to maintain a pleasant working atmosphere to make their employees stay in that company for a longer period. A few years back it was done manually but it is an era of machine learning and data analytics. Now, a company's HR department uses some data analytics tool to identify which areas to be modified to make most of its employees to stay.

The purpose is to support decisions that are based not on subjective aspects but on objective data analysis. The goal of this work is to analyse how objective factors influence employee attrition, in order to identify the main causes that contribute to a worker's decision to leave a company, and to be able to predict whether a particular employee will leave the company. After the training, the obtained model for the prediction of employees' attrition is tested on a real dataset.

Problem Statement

Employee Attrition is the gradual reduction in staff numbers that occurs as employees retire or resign and are not replaced. Employee attrition can be costly for businesses. The company loses employee productivity, and employee knowledge.

Turnover/Churn and Attrition are human resource terms that are often times confused. Employee turnover and attrition both occur when an employee leaves the company. Turnover , however, is from several different actions such as discharge, termination, resignation or abandonment. Attrition occurs when an employee retires or when the employer eliminates the position. The big

difference between the two is that when turnover occurs, the company seeks someone to replace the employee. But in the case of attrition, the employer leaves that vacancy unfilled or eliminates that job role .

Solution

When it comes to decision of whether an employee is going to stay or leave a company, his or her answer is just binomial i.e. it can be “YES” or “NO”. So, we can see our dependent variable Employee Attrition is just a categorical variable. In the case of a dependent categorical variable, we cannot use linear regression, in that case, so we have used “**LOGISTIC REGRESSION**“ as well as random forest with ANN.

So, for this project we will use evaluation of employee performance, average monthly hours at work and number of years spent in the company, among others. The dataset was spilt, using 70% for training the algorithm and 30% for testing it, achieving an accuracy of 99.1% after implementing random forest algorithm.

Literature Survey

(i) They proposed that employee turnover and attribution is explained by employees’ attitude towards job (by job satisfaction or organizational commitment or by both these attitudes). Employee turnover and attrition creates uncertainty among another workforce. The result of attrition and turnover has the emphasis of top management in nearly every organization. It indicates that the most costly and tough workforce objections facing organizations is turnover.

(ii) It utilized employee’s survey responses regarding their manager to measure people management skills. The study concludes that, the managers with the good people management skills should get rewarded from the company, which would lead to the decrease in employee turnover. Ideally, better people managers should get perks and promotions, which would help with ensuring company’s success. Also, consistent feedback regarding employees is required, which keeps the employees engaged and accountable towards their work. The study also analyzed that job which requires low skills (for e.g.

Truckers, etc.) is much easier to perform as compared to the jobs which require highly skilled workers.

(iii) This paper interprets that organizations should pay employees depending on their performances, along with various incentives such as individual bonus, lump sum bonus, profits sharing and other benefits. So, if these are put in place they will decrease employee turnover. The research also proposed that the knowledge as well as company's ideas should be shared widely among the employees too, which gives employees the motivation for staying in their present organization and become highly responsible for strong performances from the employees. Workforce optimization can also lead to the organization's success by getting work done and also establishing accountability. Good working conditions are also responsible for lower employee attrition rates. The study concluded that empowerment of employees would help with the company's success in retaining its employees.

(iv) It explored how various performance related pay (PRP) schemes impact employee turnover. The study tested whether a profit sharing has an impact on turnover as compared to other forms of PRP. The study revealed that there exists a negative relationship between aggregate measures of PRP and turnover.

Experimental Investigations

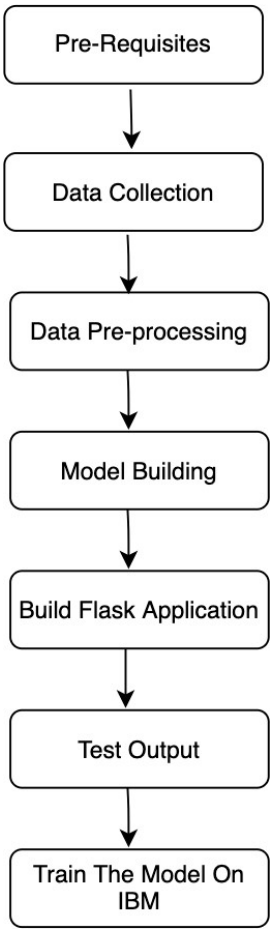
Predicting Employee Attrition is determining whether an employee will be resigning from a company or not. To achieve this result, the company will look at the following information for each employee:

- Satisfaction level with the company
- Last Evaluation Score
- Number of Projects completed
- Average Monthly Hours
- Years employed in the company
- Any accidents at work
- Promotion in the last five years
- Employee's department
- Employee's salary bracket

Looking at all this information filled by an employee, the algorithm calculates looking at these statistics whether an employee will resign or not. After the calculation, the model infers whether the employee will continue working with the company or not.

From our experiment we noticed a few trends in the relationship between the employee attributes and whether the employee leaves or not. These were our findings: Maximum number of employees belong to the technical department and employees from this department tend to leave. Employees with a lower salary are more likely to leave as compared to those with higher salaries.

Flowchart



Hardware Software Specifications

Hardware

- 8 GB RAM
- Internet Connection

Software

- Python IDE
- Anaconda
- Jupyter Notebook
- Spyder
- IBM Watson Studio
- IBM Cloud
- HTML/CSS

Conclusion

In this project, we have implemented “Predicting employee attrition”. The most significant drawback of existing organization’s data models and database is that, they contain lots of redundant data and predicting something with precision is quite challenging. We implemented a precision model for predicting employee attrition using ANN which is an artificial intelligence technique.

Using the ANN approach with random forest and logistic regression was accurate for this project as it resulted in high accuracy and low running times with efficient memory utilization. The model presented in this paper has an accuracy that almost touches 100%. It is a highly robust and scalable technique to handle all sorts of noise from huge datasets and convert the data into a ready acceptable form for precision results.

Future Scope

In the near future, we like to implement the proposed model in real-world organizations to enable them to learn the reasons for employee turnover. The research would go in the direction to make this model a “Predictive Mode” and solve various issues, i.e. Advanced ones not predicting- “Who is going to Leave?” but also “Why the Employees are doing turnover?”. The model will

become more accurate, scalable and ready to implement as such in top IT organizations HR departments.

Bibliography

[1] Ozolina-Ozola, I. Reducing Employee Turnover in Small Business.

[2] Hoffman, M., & Tadelis, S. (2018). People Management Skills, Employee Attrition, and Manager Rewards: An Empirical Analysis

[3] Ongori, H. (2007). A review of the literature on employee turnover.

[4] O'Halloran, P.L. (2012). Performance pay and employee turnover

Journal of Economic Studies, 39(6), 653–674.

https://www.knime.com/blog/predicting-employee-attrition-with-machine-learning

http://ijiset.com/vol7/v7s9/IJISET_V7_I9_07.pdf

https://www.analyticsvidhya.com/blog/2020/11/employee-attrition-analysis-using-logistic-regression-with-r/