

PREDICTING EMPLOYEE ATTRITION USING NEURAL NETWORK

Yaswanthan A*, Sridhuresh S, MSc(Decision and Computing Science) (5 years Integrated Course), Coimbatore Institute of Technology, Coimbatore.

Abstract— HR attrition is a major problem for any industries (manufacturing or services), retaining key talent and identifying the people who need to be improved was tough task. Losing top key talent and filling them with equivalent talent was paramount challenge and would cause organization a huge loss in terms of time and cost.

To provide solution to this problem, we propose a neural network model to predict the person who most likely to move out of the organization from the set of employee data. The employee data consist of lot of independent variables and decision variable is attrition (Yes/No). With machine learning techniques such as RANDOM FOREST, Neural Network model would be built, this model would be a handy support tool to the HR manager to identify the person who likely to move out of the organization. Based on the impact and insight from this predictive model, insights can be drawn and implement actions to improve the mindset of the person in question. This paper would discuss about the data, data classification, neural network model such as RANDOM Forest, accuracy of the model built and insights drawn based on study.

Keywords: Neural Network, Random forest, predictive model, HR Insights

INTRODUCTION: Companies in India as well as in other countries face challenges of recruiting and retaining talents. Losing top key talent and filling them with equivalent talent was paramount challenge and would cause organization a huge loss in terms of time and cost. With various offers from many companies have the potential to transform the talented employees, which brings out the importance of predicting employee attrition.

PROPOSED ALGORITHM: We have used the data from great learning, “Introduction to Machine learning” which was prepared for workshops. The data consists of many employee details. We have undertaken employee age, monthly income, working year in the company, environmental satisfaction, over time, job involvement and employee attrition as parameters. The following input features were extracted from the dataset:

1. Employee attrition-the data where employee have the opinion (0 or 1)
2. Monthly Income-salary of the employee
3. Working year in the company-total experiences in the

company

4. Total working hour-how many hours the employee works in a day

5. Age-employee's age

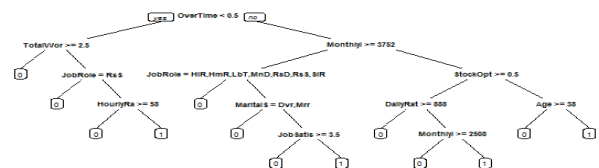
6. Environment Satisfaction-the satisfaction of the employee according to the environment

7. Job Involvement- these says the involvement of the employee.

8. Over time-does the employee accepts the over time working hours.

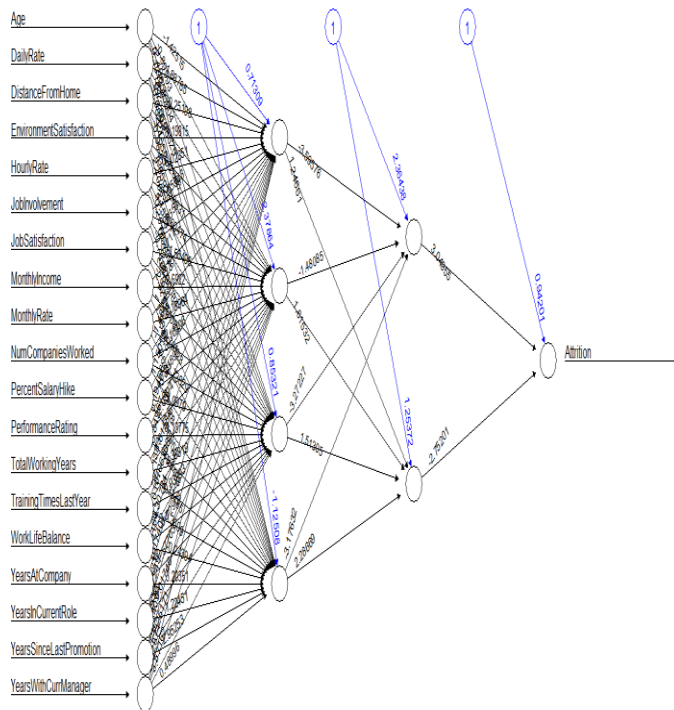
In this data, there are two types of variable. One is dependent variable and another is independent variable. Some of the independent variables are monthly income, working year in the company, total working hour, age, Environment satisfaction, job involvement, over time. The dependent variables are employee attrition. The negative value indicating the employee will drop out and positive value indicates that the employee will continue in the company. We predict whether the employee is going out of the company or not. Our algorithm pinpoints the employee is predicted to drop-out, which allows the human resources manager and his team to take necessary steps to prevent or reduce employee attrition.

METHODOLOGY: The data was analyzed to find which factor has maximum influences towards attrition. The influenced attributes were taken to construct the artificial neural network. The data was divided into two sets (test data, train data) which are useful to create a neural network model. Due to the large amount of data set, random forest algorithm was employed. The neural network model was created trained using trained data set; the neural network was validated using the test data set. For the purpose of cross checking the neural network, Root Mean Square Error (RMSE) method and K-fold cross validation method has been done.



NEURAL NETWORK:

Artificial neural networks are suitable to predict the employee attrition as there are a large number of inputs. We construct an artificial neural network with the input layer: employee attrition, Environment Satisfaction, monthly income, working year in the company, total working hour, hourly Rate, age. Output layer consists of single node predicting whether employee is going to drop-out. We add a hidden layer of 2 neurons in the neural network between the input and output layer. The number of neurons in the hidden layer was experimentally determined to get best possible results. The diagram shows the structure of the neural network used to predict employee attrition. It gave best results in our experiments.



CROSS VALIDATION OF NEURAL NETWORK: In predicting employee attrition, our focus is to capture all employees who are going to drop-out and thus, minimizing the error is very important. Thus, we can use MSE or cross validation method of Neural Network.

In MSE we can predict the error of the neural network for the trained data. When we introduce new data it cannot find the error of the neural network. There is the possibility that the artificial neural network predicts wrong status of the employee in the company, which tends the company to loss. For the process we have split the data, which was the simplest form of the cross validation of the network which is known as the hold out method. We have also used the k-fold cross validation of networks. We compute the RMSE value for each training set, that are stored in the matrix.

RESULTS AND DISCUSSIONS: This algorithm predicts the employee attrition rate in MNC's. When the attained output value lie in negative, the employee is likely to

drop out from the company and when the output is in positive, the employee is likely to continue in the MNC's. Overtime is the most important factor contributing towards attrition. Overtime remains the single most important factor contributing towards attrition both for the CART and Random Forest models. Young people with less year at the company contribute to the attrition

CONCLUSION: We propose an algorithm to predict employee attrition using an artificial neural network. This method pin-point the employees who are wish to drop-out. Apart from MNC's, the proposed algorithm can also used in many developing company, which are becoming increasingly popular in recent years. It also summarizes the data and finds which attribute has maximum percentage of influences of the employee's attrition.