**Assignment 2**

**Topic - Exploring the factors affecting Attrition.**

**Presented By :**

**Shubham Kumar**

**Roll No. D18038**

**Problem Statement :**

Let’s start with a term called organization, So what is an organization? – it is group to people working under single leader and working to achieve same objective. But , an organization also called as business usually faces a common problem of employees leaving the job for better growth, no job satisfaction , no proper hike in salary etc. There could be N – number of reason which can contribute to Attrition of employee.

Listing down some the problems which organization faces due to Attrition –

* Money invested on training of employee is not wasted.
* Building a backup resource is time taking task if any senior guy leaves in the middle of any project.
* Following a new hiring process again requires money.
* If New hired employee ,leaves company within a year.

Let me tell you incident of one my good friend, who joined one of the top MNC . He was complete blend of smartness in terms of work, focused , go – getter attitude , but the problem which every IT guy faces if you perform good then it is expected that you have to the capability of handling more work all alone and that too without additional perks provided . That what happened with him, he use to devote more than usual 9 working hours . Though he was well appreciated by every single manager with he worked but if a person not getting rewarded in money terms then there is no point working anything extra. So , having experience and sound knowledge of the domain on which he was working on he left the company joined another one. MNC has lost, one of its valuable resource though it is different some else will step in his shoes , but the money what company has invested on a person by giving him training sessions is now wasted , this learning he will be using for some other organization in which he will be moving in.

Though this is one of the many reasons which I wanted to highlight with the help of above incident . As in the case from above incident , the person was not satisfied with salary hike or hourly salary . Likewise , there can be several other reasons like – if a person is married , working late night hours doing overtime tends to hamper there work life balance. Usually , Males who just got married (age – 28 to 32) tend to Attrite as they have responsibility of there family.

So , above we have discussed few problems which business faces due to employee Attrition and thus effecting business directly or indirectly.

**Objective :**

To find best possible set of feature among given feature in Attrition excel which can clearly define if a employee will Attrite or not . If it is ‘Yes’ , then what all preventive measures should a organization should implement to stop the employee to look for another job and remain part of organization. If there is no other option with employee except for leaving current organization then it this case , making a backup of resource well in advance so that there would not be any rush at the eleventh hour.

If it is ‘No’ , then getting insight from these set of people like what organization is doing well for these set of people (excluding all external factors like – family circumstances etc). Trying to compare with set of people who opted for ‘Yes’ .

Depending on there analysis , trying to implement new policies or changes , which can which can benefit employee as a whole.

**Analysis and Interpretation :**

**Feature : “Overtime”**

1. **Hypothesis :**

Null Hypothesis (Ho) - OverTime highly impacts Attrition rate.

Alternative Hypothesis (Ha) – OverTime has impact on Attrition rate.

1. Below is cross – tabulation performed on Attrition Vs Employee’s doing Overtime.

|  |  |  |
| --- | --- | --- |
| OverTime | No | Yes |
| Attrition |  |  |
| No | 0.9 | 0.69 |
| Yes | 0.1 | 0.31 |

1. From the above graph , we have calculated proportion of employee who are doing overtime and are more inclined towards changing job. We can say out of 100 employees 31 employees are inclined towards changing from there current organization.
2. If a person is paid Rs. 450 /day for working 9 – hours a day . Now , suppose for next onwards if we ask same person to work overtime of 3 hours i.e : 12 hours/ day , so we can see the calculation how much employee is losing.

Extra amount which was not paid to employee (Calculating for a week) = (Overtime/day \*(no of days employee did overtime))

= 3\*450

= Rs. 1350 **(So , on average employee is losing Rs. 1350 by doing overtime.) – Ho accepted.**

**Feature : “BusinessTravel”**

1. **Hypothesis :**

Null Hypothesis (Ho) - BusinessTravel highly impacts Attrition rate.

Alternative Hypothesis (Ha) – BusinessTravel has impact on Attrition rate.

1. Below is cross – tabulation performed on Attrition Vs Employee’s doing Overtime.

|  |  |  |  |
| --- | --- | --- | --- |
| BusinessTravel | Non-Travel | Travel\_Frequently | Travel\_Rarely |
| Attrition |  |  |  |
| No | 0.92 | 0.75 | 0.85 |
| Yes | 0.08 | 0.25 | 0.15 |

1. From above graph, we have calculated proportion of people who travel frequently and inclined to changing there organization. We can say about 1/4th of the employee who travel frequently tend to change there current job.
2. Employee who have field job are required to travel frequently , generally people prefer Non- travel or Rarely Travel jobs as we can see from the above cross – tabular table . Generally , it happens that job which require frequent travel example a Network engineer in any Telecom company won’t be preferred by people with age group(30 – 40 years) , reason could be age and employee want to have well settled life. **(Ho is accepted)**

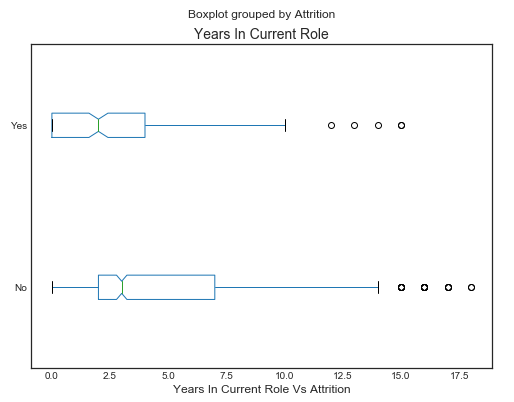
**Feature : “YearsInCurrentRole”**

1. **Hypothesis :**

Null Hypothesis (Ho) – YearsInCurrentRole highly impacts Attrition rate.

Alternative Hypothesis (Ha) – YearsInCurrentRole has impact on Attrition rate.

1. Below is box plot drawn between Attrition Vs YearsInCurrentRole.



1. From the above Box plot , employee with exp. of 2.5 years in their currentJobRole tend to seek for promotion , looking at 2nd quartile values for employee who Attrite ‘Yes’ is less than 2nd quartile values for employee who Attrite ‘No’.
2. Considering median value , employee who Attrite ‘Yes’ . Employee with exp. of 2.5 years in their currentJobRole tend to get bored this can be either due to the large gap in last promotion they had , or they want to take higher job responsibilities.

From Heatmap plot we have calculated features that have higher correlation. **(Ho is accepted)**

**Feature : “MaritalStatus”**

1. **Hypothesis :**

Null Hypothesis (Ho) - MaritalStatus highly impacts Attrition rate.

Alternative Hypothesis (Ha) – MaritalStatus has impact on Attrition rate.

1. Below is cross – tabulation performed on Attrition Vs MaritalStatus.

|  |  |  |  |
| --- | --- | --- | --- |
| **MaritalStatus** | **Divorced** | **Married** | **Single** |
| **Attrition** |  |  |  |
| **No** | 0.9 | 0.88 | 0.74 |
| **Yes** | 0.1 | 0.12 | 0.26 |

1. From above cross – tabular , we have calculated proportion of employee who are ‘Single’ and have attrition ‘Yes’ . Out of 100 single employee 26 single employee tend to move to different organization.
2. Employee with MaritalStatus as ‘Single’ , tend to switch from one organization to another organization . They are completely independent and have no responsibility on anyone .

**Fit a Decision Tree Classifier :**

What we did , performed Exploratory Data analysis among difference features present in Attrition excel . First checking for the any Null value in the dataframe so as to have any prior knowledge as this might not cause any issue in selecting appropriate feature during selection.

In case of Categorical variable , we have taken feature like – Department , BusinessTravel, EducationalField , Gender, JobRole , MaritalStatus , OverTime . Now , we draw a cross- tab between Attrition Vs above listed Categorical variable in order to find best possible which is actually impacting the Attrition rate in any organization.

Now , in case of Numerical variable , we tried to represent the data with the help of boxplot which are grouped by ‘Attrition’ w.r.t CurrentJobRole . Though we have not taken this variable in consideration in feature selection , as per Heatmap plot there is correlation between CurrentJobRole and feature like TotalYearsOfExp that are highly correlated , hence tried to reduce the number of features depending on correlation coefficient values obtained from the Heatmap plot. This is shown in the Python file as well.

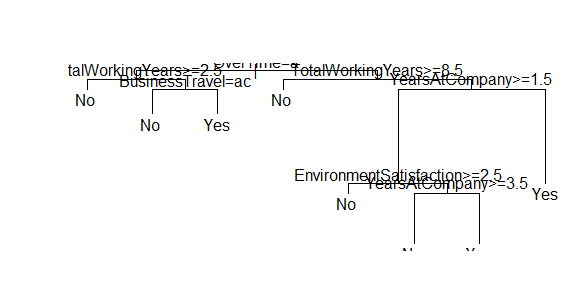
Now , No model can fit categorical variable values , hence we first converted categorical variable using Pandas libraries dummy variable which act as a feed to ML model.

Feeding the model with train data set in order to train our model to get the insight of pattern followed and then exposing our model to test data set to for checking our efficiency/ accuracy of the trained model.

We drew confusion matrix just to calculate how much we are close to the expected value i.e how much our predicted and expected values differ from each other . More the number of counts matching the expected value with predicted value , more will our accuracy of the model.

Below Plot is obtained from R – Studio though it is not clear , attaching below code for getting actual plot of Decision Tree :





**Suggested Actions :**

* Optimizing recruitment by verifying number of companies an employee has worked on and how much time he has spent / work exp. in each organization.
* Enhancing proper recognition and reward program for employee based on there performances , giving bonus from there variable pay as an extra perk for there good work.
* Improving work culture so that employee should have work life balance. Engaging employees in various cultural and social events in office so that they should not get bored in the work schedule and remained stressed out.
* Provide better facilities to both Male and Female – like if Females are given free cab facility if working late night hours then Male who don’t have there own means of transport should be allowed to use that cab facility.
* Offer continuous promotion if employee is performing good along with amount of experience he/she is having in their currentJobRole.
* Cross – skilling employee’s so that they should not get bored with doing same kind of stuff doing repeatedly.

**Python Code :**

