Name -ROHIT SHARMA

Internship Program – Data Science With Machine Learning And Python

Batch – Jan 2022 – Mar 2022

Certificate Code -TCRIB2R133

Shape, arrow

Description automatically generated

Technical Coding Research Innovation, Navi Mumbai,

Maharashtra, India-410206

(HR Employee Attrition Data Analysis)

A Case-Study Submitted for the requirement of

**Technical Coding Research Innovation**

For the Internship Project work done during

**DATA SCIENCE WITH MACHINE LEARNING AND PYTHON INTERNSHIP PROGRAM**

by

**ROHIT SHARMA** (TCRIB2R133)

Rutuja Doiphode

CO-FOUNDER &CEO

**HR ATTRITION**

Human Resource analytics deals with the analysis of the HR department in the organization. The goals of HR analytics are improving employee performance and getting better returns on investment.

Attrition in human resources refers to the gradual loss of employees overtime. Losing employees is not good for companies. Losing employees will cost more for the company because it will eventually involve hiring new employees and training them again. HR professionals have a responsibility to help the organization by maintaining a good work culture, good work environment etc.

In this assignment, we need to predict whether a given employee will leave the organization or not.

We will create a model with the following steps:

● Import the relevant packages

● Download and explore the dataset

● Perform EDA, Apply dataset for preprocessing

● Predict the target columns

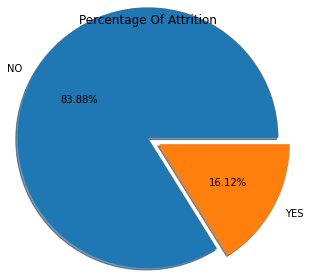
What are the reasons or features on which attrition depends (features we have in the dataset

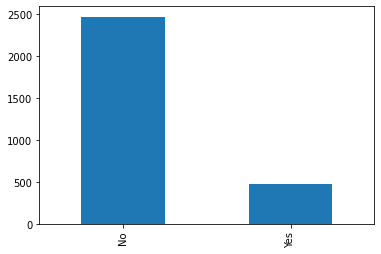
**Age, Business travel, Department, Distance from home, Education, Environment satisfaction, Gender, Hourly Rate, Job Involvement, Job Level, Job Role, Job Satisfaction, Marital Status, Monthly Income, number of companies worked**



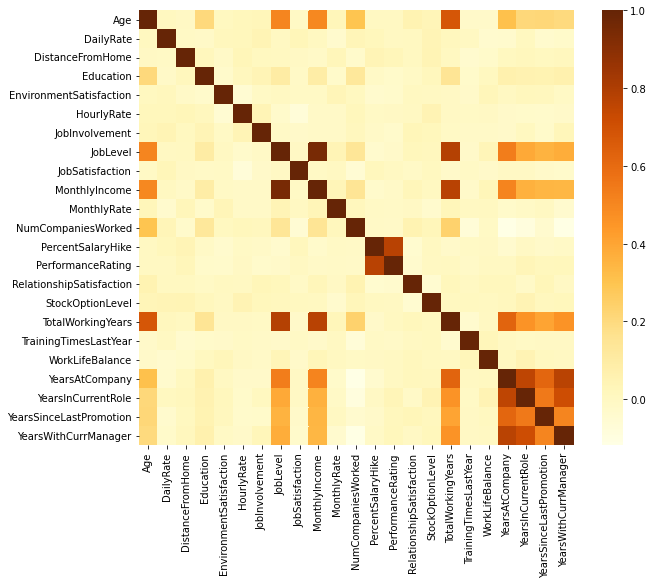
**Data Analysis –**

We will see the overall attrition percentage from the datasets.





From the above matrix we can find features that affect attrition (target) features. we can some features are related like –

● Age and Total Working Years / Monthly Income / Job Level / Num of companies Worked

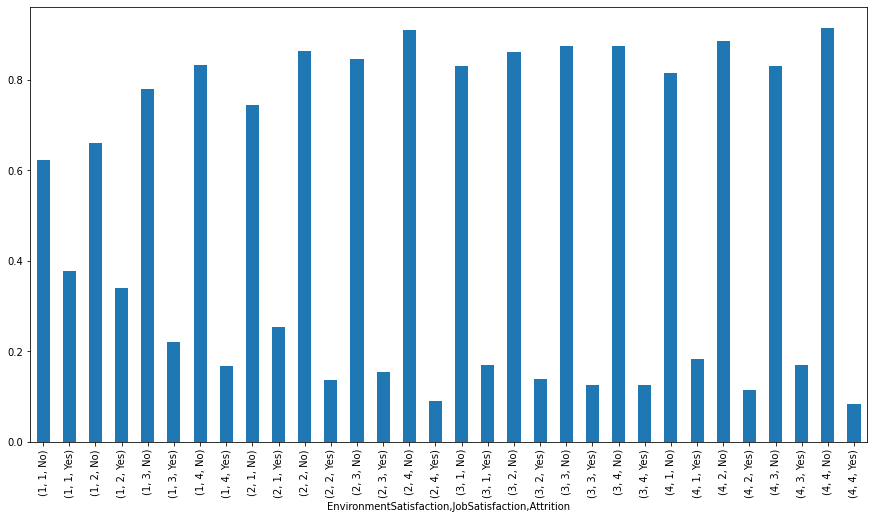
● Job Level and Total working years

● Job Level and Years at company

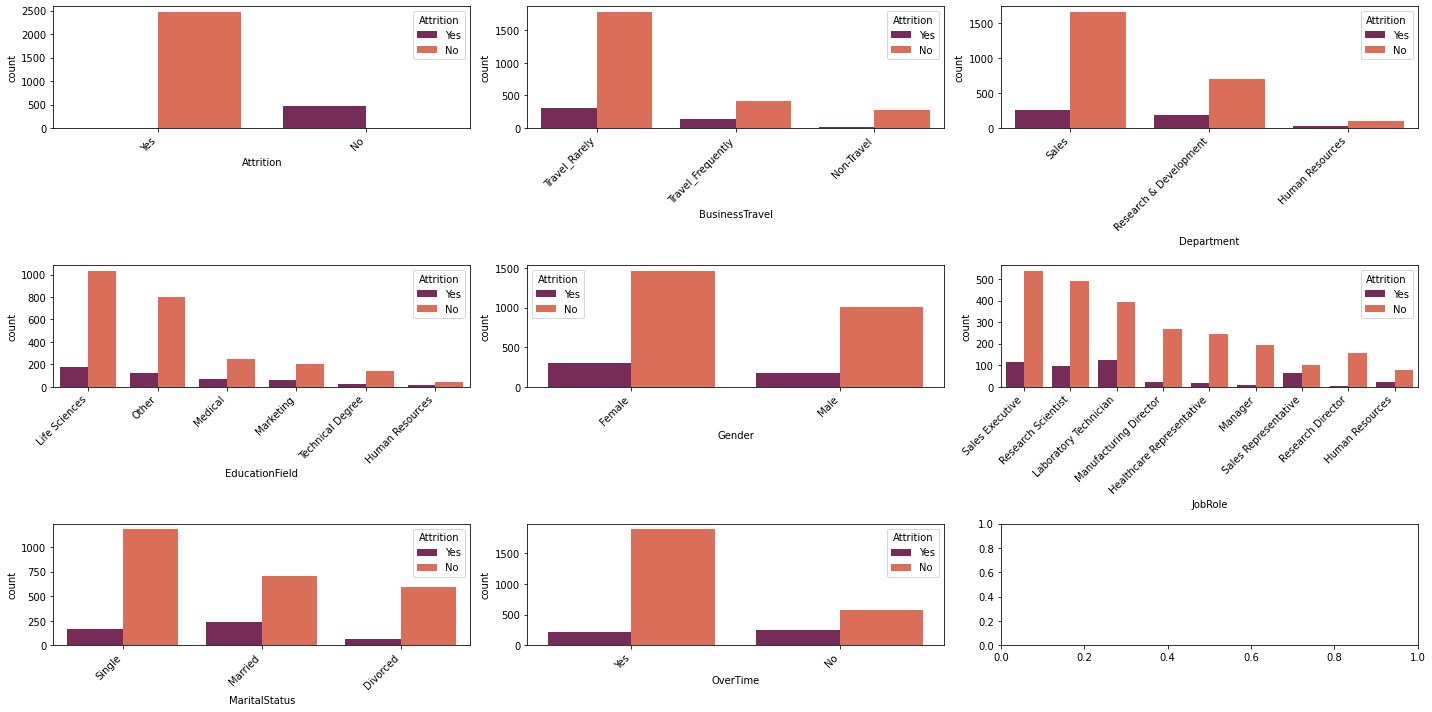
● Income and Working years

● Job Level and Monthly income

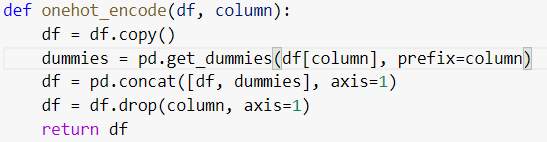
**Attrition VS Environment Satisfaction and Job Satisfaction -**

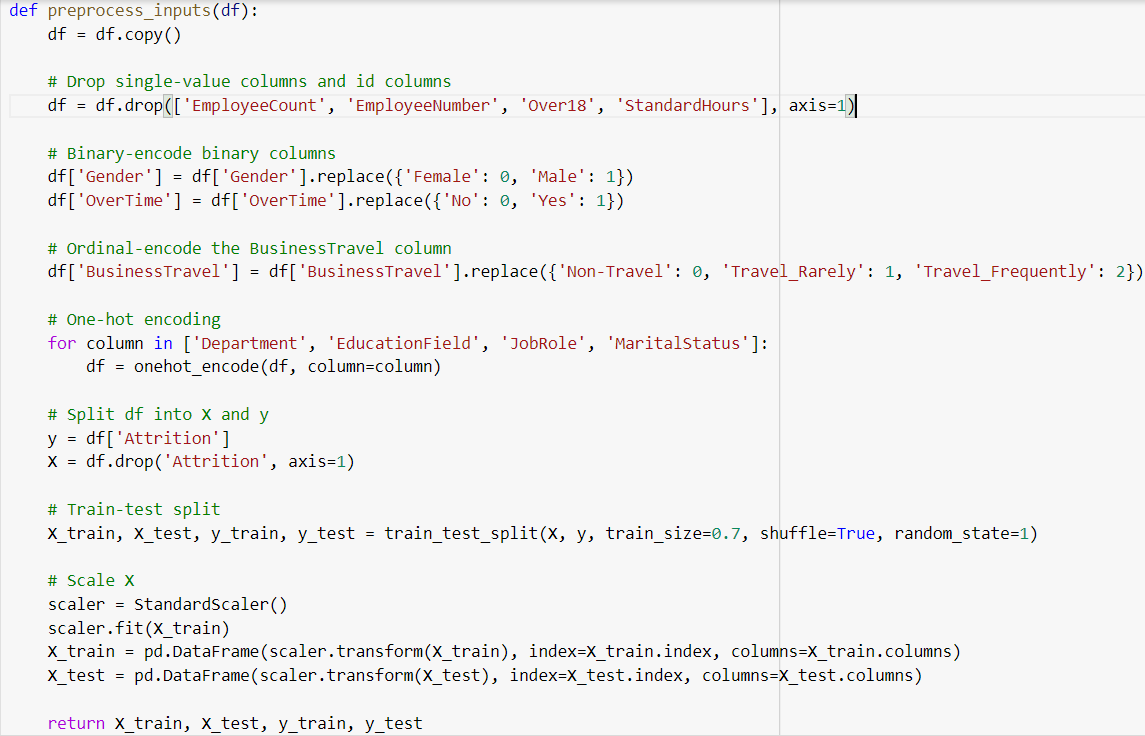


As we can see, attrition level decreases as environment satisfaction level and job satisfaction level increases



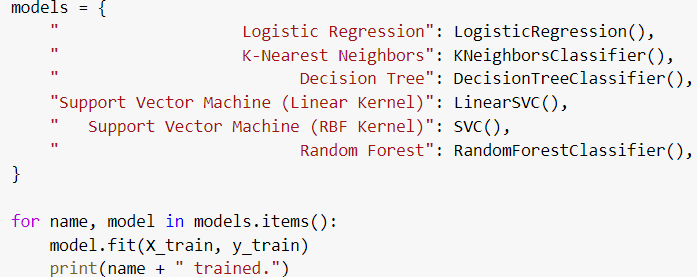
**PREPROCESSING**





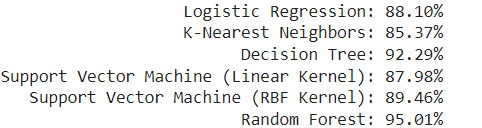


**Splitting dataset for training and testing.**





RESULT



**CONCLUSION**

For classification I have compared classification algorithms such as Logistic regression, KNN, Decision tree, SVM, Random forest.

I got best result by using Random Forest model with the accuracy of  ***95.01%***.