**PROFESSIONAL TRAINING REPORT**

**At**

**Sathyabama Institute of Science and Technology**

# (DEEMED TO BE UNIVERSITY)

Submitted in partial fulfillment of the requirements for the award of Bachelor of Engineering Degree in Computer Science and Engineering

By

**Peravali Deepesh**

**(Reg.no. 39110770)**



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**SCHOOL OF COMPUTING**

**SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY**

**JEPPIAAR NAGAR, RAJIV GANDHI SALAI,**

**CHENNAI-600119, TAMILNADU.**

**APRIL 2022**

SATHYABAMA

## INSTITUTE OF SCIENCE AND TECHNOLOGY

## (DEEMED TO BE UNIVERSITY)

**Accredited with Grade “A” by NAAC**

**JEPPIAAR NAGAR, RAJIV GANDHI SALAI, CHENNAI -600119**

[**www.sathyabama.ac.in**](http://www.sathyabama.ac.in/)

## SCHOOL OF COMPUTING

**BONAFIDE CERTIFICATE**

This is to certify that this Professional Training Report is the bonafide work

of **Peravali Deepesh (Reg.no.39110770)** who carried out the Project

entitled **“Attrition Analysis and Prediction in an IT Organization”** under our supervision from **February 2022** to April 2022.

**Internal Guide**

**Mr.Rajasekar.P**

**Head of the Department**

**Dr. S.VIGNESHWARI, M.E., Ph.D.,**

**Dr.L.LAKSHMANAN, M.E., Ph.D.,**

**submitted for Viva voce Examination held on**

**Internal Examine External Examiner**

# DECLARATION

**I, Peravali Deepesh (Reg.no.39110770)** hereby declare that the Professional Training Report on **“Attrition Analysis and Prediction in an IT Organization”** done under the guidance of **Mr.Rajasekar.P** at **SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY** is submitted in partial fulfillment of the requirements for the award of Bachelor of Engineering degree in Computer Science and Engineering.

**DATE: 10-04-2022 Peravali Deepesh**

**PLACE: CHENNAI SIGNATURE OF THE CANDIDATE**

# ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to **Board of management** of **SATHYABAMA** for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

I convey my thanks to **Dr. SANKARAN.V, M.E., Ph.D., Manager of our HCL Internship** for providing us the necessary support and details at the right time during the progressive reviews.

I convey my thanks to **Dr. T.Sasikala, M.E., Ph.D., Dean of the Department, Department of Computer Science and Engineering** for providing us the necessary support and details at the right time during the progressive reviews.

I convey my thanks to **Dr. S. VIGNESHWARI, M.E., Ph.D.,and Dr.L.LAKSHMANAN, M.E., Ph.D.,Heads of the Department, Department of Computer Science and Engineering** for providing us the necessary support and details at the right time during the progressive reviews.

I would like to express my sincere and deep sense of gratitude to my Project Guide **Mr.Rajasekar.P** for his valuable guidance, suggestions and constant encouragement paved way for the successful completion of my project work.

I wish to express my thanks to all Teaching and Non-teaching staff members of the Department of **COMPUTER SCIENCE AND ENGINEERING** who were helpful in many ways for the completion of the project.

# CERTIFICATE

**ABSTRACT**

The study seeks a distinctive and efficient Machine Learning System for the prediction of Employee Attrition. In any industry, employee attrition is a big problem, whether it is about employee attrition of an organization or customer of an e-commerce site. If we can accurately predict which customer or employee will leave their current company or organization, then it will save much time, effort, and cost of the employer and help them to hire or acquire substitutes in advance, and it would not create a problem in the ongoing progress of an organization. In this chapter, a comparative analysis between various machine learning approaches such as SVM, Decision Tree, KNN, Random Forest, Logistic Regression,Gradient Boosting and Nureal Network is presented. The presented result will help us in identifying the behaviour of employees who can be arrired over the next time. Experimental results reveal that the logistic regression approach can reach up to 85% accuracy over other machine learning approaches

# TABLE OF CONTENTS

**CHAPTER TITLE PAGE NUMBER**

**NO**

|  |  |  |
| --- | --- | --- |
|  | **ABSTRACT** | **VI** |
| **1.** | **CHAPTER1-INTRODUCTION** | **09** |
|  | 1.1 Aim of Study | 11 |
|  | 1.2. Limitation of study | 11 |
|  | 1.3. Problem Statement | 12 |
|  | 1.4. Methodology | 12 |
|  | 1.5. The Study Region and Data | 13 |
| **2.** | **CHAPTER 2 -LITERATURE REVIEW** | **14** |
| **3.** | **CHAPTER 3** | 15 |
|  | 3.1 Design and Architecture | 15 |
|  | 3.2 Technologies Used In The Proposed System | 16 |
|  | 3.2.1 Machine Learning | 16 |
|  | 3.2.2 Logistic Regression | 16 |
|  | 3.2.3 KNN | 16 |
|  | 3.2.4 Decision Tree | 17 |
|  | 3.2.5 Random Forest | 17 |
|  | 3.3 Data Analysis | 19 |
|  | 3.4 Result and Evaluation | 20 |
| **4** | **CHAPTER 4 - DESIGN** | **25** |
|  | UML DIAGRAMS | 25 |
|  | 4.1. Use Case Diagram | 26 |
|  | 4.2. Sequence Diagram | 27 |
|  | 4.3. Activity Diagram | 28 |
|  | 4.4. System Diagram | 29 |
| **5** | **CHAPTER 5-SCREENSHOTS** | **30** |
|  | 5.1 ML CODE SCREENSHOTS | 30 |
|  | **CONCLUSIONS…….** | **39** |
|  | **REFERENCES….**  **GITHUB LINK** | **40** |
|  |  |  |

**Attrition Analysis and Prediction in an IT Organisation**

**CHAPTER1**

**Introduction**

Today attrition is one of the major problems faced by industry across the world. It is the most burning issue for the industry, and high attrition rates lead to many issues in the boundary of the organization like losing the talents and knowledge, cost related to training and administration, and recruitment. It is observed that many attributes lead to the attrition of an employee. Which includes working environment, job satisfaction, employer’s behavior, job timing, and most important is salary or incentives. Also, the prediction model plays an essential role in finding the behavior of employees.

Timely delivery of any service or product is the primary goal of any organization in recent days due to high competition in industries. If a talented employee leaves unexpectedly, the company is not able to complete the task at defined times. It may become the reason for the loss of that company. Therefore, companies are interested in knowing the employee’s attrition. They can make a proper substitute or arrangements earlier.

There may be various reasons for employee attrition, which include less salary, job satisfaction, personal reasons, or environmental issues if the employer terminates an employee for any reason. It is known as involuntary attrition. On the other hand, voluntary attrition is known as the left of an employee by their side. This kind of attrition is a loss for the company if he or she is a talented employee. In the present scenario, everyone wants a higher salary and job security. Therefore, employees leave jobs immediately if they got a better chance in other places.

In the recent era of computer science, machine learning approaches play an important role in employee attrition prediction. These approaches provide predictions based on historical information of the employee, such as age, experience, education, last promotion, and so on. Based on the prediction results HR department have prior knowledge about employee attrition. The HR department also has preplanned recruiting employees as a substitute for the employee who is interested in leaving in the coming days.

Various researches have also studied the performance of different machine learning approaches have discussed various reasons or factors that are involved in employee attrition. They have also investigated that talented employee replacement is a time-consuming and challenging task. It is also a significant factor in loss in business.

Compensation is one solution to decreasing the attrition rate. Moncaarz et al. (Moncarz et al., 2009) have discussed how attrition can be decreased by providing better compensation.

Punnoose and Ajit (Ajit, 2016) have provided a comparative analysis of various machine learning approaches for employee turnover. Tree-based approaches are also used to predict employee attrition (Alao & Adeyemo, 2013). Jantan et al. (Jantan et al., 2010) have compared tree-based methods with other traditional machine learning approaches. Radaideh and Nagi (Al-Radaideh & Al Nagi, 2012) uses the decision tree for employee attrition prediction. In their work, they have found that job title is an es-sential feature of attrition, whereas age is not a very important feature.

Saradhi (Saradhi & Palshikar, 2011) uses various machine learning approaches for employee attrition prediction. They have taken a database of 1575 records with 25 features of employee and applied various classification approaches to predict attrition. They have shown that SVM has higher accuracy, which is 84.12%.

Due to confidentiality and noisy HR data, sometimes prediction has higher accuracy. It is difficult to generalized predictions for different organizations and employee roles.

Previous studies presented accuracy as a primary evaluation standard for attrition prediction. Various machine learning approaches are used and evaluated in different datasets. It is challenging to conclude that which model is best for attrition prediction. The rate of employee attrition is always less than the employee who stays in the organization. Therefore, datasets are always imbalanced. Accuracy measures are not reliable for imbalanced datasets. So that it is desired to have an accurate model to enhance the prediction accuracy of the models. Which provides better results to employers. Based on the accurate prediction results employers and HR department know the behavior of their employees

.

* 1. **Aim Of Study**

The aim of this chapter is to provide a comparative analysis of different machine learning approaches for analysis and employee attrition prediction. Here we have significantly enhanced the training process to solve the imbalanced class problem**.**

* 1. **Limitations Of Study**

The Scope of the research is divided into five parts - First part contains the study of the factors influencing employee attrition. Second part has study on EDA that mean finding the missing values reading the data,displaying the data and correlation between attritbutes. Third part has the study on Data Preprocessing. Fourth part has Model Evaluation. In fifth part Model Deployment.

The budgeting of the employee turnover to indicate how much the company has saved the

employee turnover budget after retaining the valuable employee is not included in in the

Dissertation as it gets into the completely other filed of human resource management budget. And hence, I have limited my research work till improvisation of employee retention. Limitation includes implementation of Prediction System for Employee Attrition due to less training data set, implementation of Decision-Making System for Retention due to complex development of the Classification Tree and Algorithm and inconsistent data, accuracy of Decision result, limited access to employee dataset due to GDPR. My dissertation has a major contribution in the field of HR Analytics in improving the accuracy

of the prediction of employee attrition and advancing the application for helping the HR and Project Managers to improve the retention rate of valuable employee by building Decision tree for selecting the valuable employee and finding the factors influencing them to resign, thereby saving the employee turnover budget of the company.

* 1. **Problem Statement**

Employees are the most important part of an organization. Successful employees meet deadlines, make sales, and build the brand through positive customer interactions.

Employee attrition is a major cost to an organization and predicting such attritions is the most important requirement of the Human Resources department in many organizations. In this problem, your task is to predict the attrition rate of employees of an organization.

* 1. **Methodology**

This paper discusses supervised learning methods of classification, since we know of the existence of two classes working and left. This section outlines the theory behind each machine learning algorithm

1. Logistic regression
2. SVM
3. KNN
4. Decision Tree
5. Nureal Network
6. Random Forest
7. Gradient Boosting

The proposed work analyses the respective dataset to detect the most influential

features that affect the prediction and builds a predictive model according to the following phases.

1. Gathering employees’ data: IBM dataset [16] has been used.

2. Elimentary Data Analysis(EDA):checking the missing values,finding the correlation between attrition and other attributes.

3. Data Preprocessing:using Encoding techniques converting categorical data into numerical data and splitting it into training and testing data

4. Model Evaluation:Fitting the trained data into several machine learning algorithms and gets the model trained.

5. Model Deployement:Finding the accuracy score from the models.

# The Study region and data

The metrological data including salary, business travel, daily rate, over time, Employee Satisifaction, Last Pramotion, Salary hikes, Age, Department, Distance from Home, Hourly Rate, Job role, Number of companies worked etc. Although, it has all attributes to predict the employee attrition.

**CHAPTER 2**

**LITERATURE REVIEW**

Employee attrition refers to the gradual loss of employees over time. Most literature on employee attrition categorizes it as either voluntary or involuntary. Involuntary attrition is thought of as the mistake of the employee, and refers to the organization firing the employee for various reasons. Voluntary attrition is when the employee leaves the organization by his own will. This paper focuses on voluntary attrition. A meta-analytic

review of voluntary attrition found that the strongest predictors of voluntary attrition included age, pay, and job satisfaction. Other studies showed that several other features,

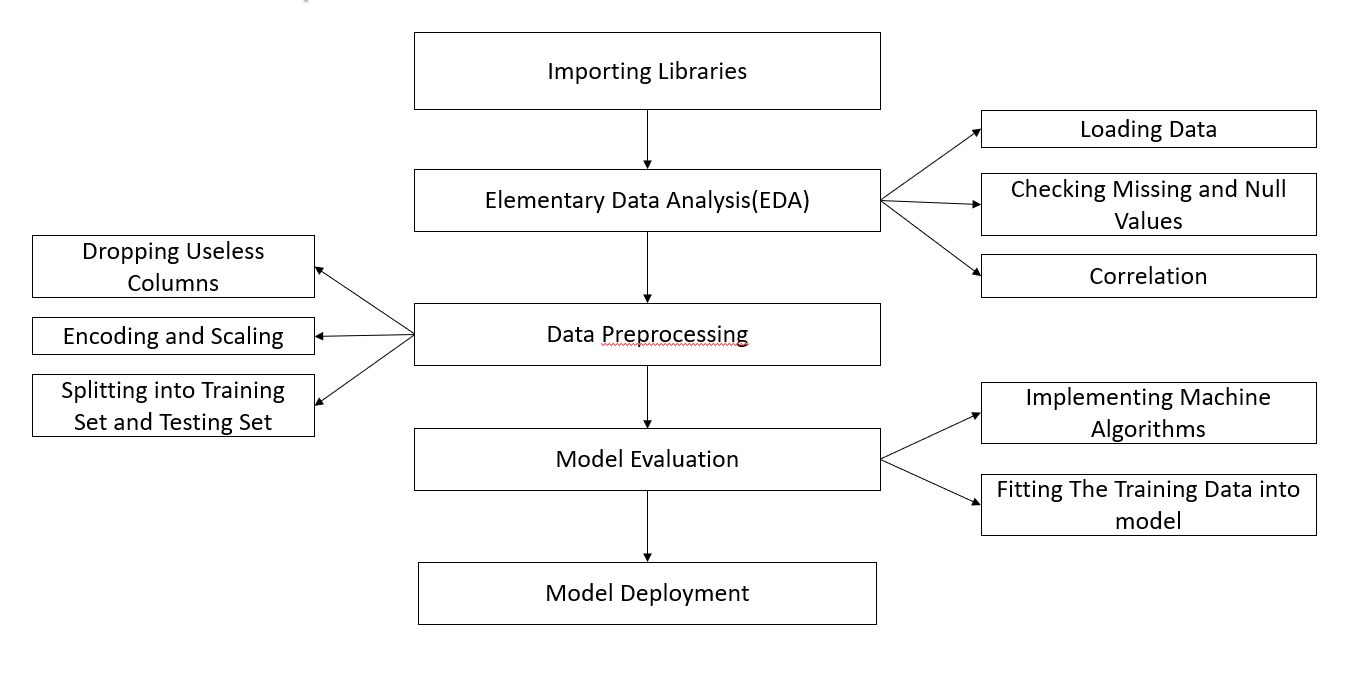
such as working conditions, job satisfaction, and growth potential also contributed to voluntary attrition.

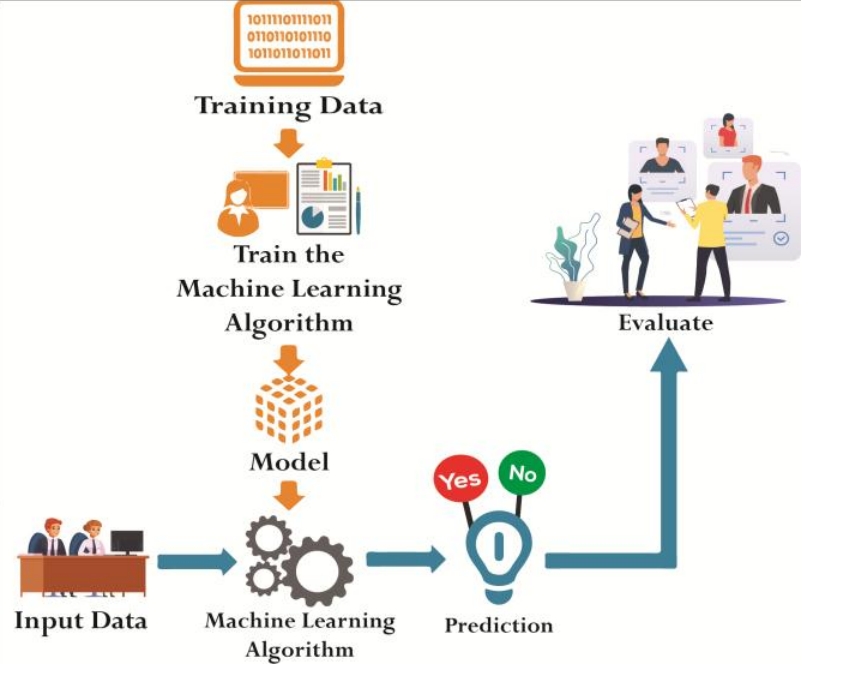
Organizations try to prevent employee attrition by using machine learning algorithms to predict the risk of an employee leaving, and then take pro-active steps for preventing such an incident. Employee Attrition is mainly the normal flow of people out of an organization, due to career or job change, relocation, illness and so on. Employee Attrition is the percentage of employees leaving the organization for what so ever reasons. Employees can leave the organization for many personal as well as professional reasons. So basically there are two types of turnover, one is voluntary turnover which is decided by the employee, and the other type of turnover is decided by the company and that is why it is called involuntary turnover. Involuntary turnover generally happens when performance of the employee is not up to the expectations. Retention is also necessary for the growth and stability of an organization. The high attrition rate causes when there are more employment opportunities in the market. Currently the employee attrition is one of the major issue faced by HR managers. There are so many working employees who are

not satisfied due to one of the aspect which is not fulfilled by the organization which results in higher attrition rate.

**CHAPTER 3**

**3.1 DESIGN AND ARCHITECTURE**





The proposed system consists of different machine learning algorithms. To build model, we take employee dataset which includes all past and present records of the employees, then we perform data preprocessing (Data Preprocessing is that step in which the data gets transformed, or encoded, to bring it to such a state that the machine can easily analyze it). We have divided dataset into two parts one is train data and second one is test data. Most of the data is used for training and smaller portion of data is used for testing (Train: 70%,Test: 30%). The aim of training is to make a prediction correctly as often as possible. The test data is used to see how well the machine can predict new answers and to validate machine learning model behavior. Afterward, using different machine learning algorithms wehave build the model. After building model, user can give the new input data to the system. Furthermore, user can choose algorithm according to their choice and check the result. Output of the system is in two forms - one is graphical representation and other is in polar form that is ‘Yes’ or ‘No’ format. After evaluating result the reason behind the attrition is also given by the system.

# 3.2 TECHNOLOGIES USED IN THE PROPOSED SYSTEM

# 3.2.1 MACHINE LEARNING

Machine Learning is most important technology towards data analysis for quality prediction and evaluation. There are various algorithms in machine learning which are used to predict the appropriate class of new or unseen data. In our system we used different machine learning algorithms to find out the reasons for employee attrition. The machine

learning algorithms which are used in system are described below:

**3.2.2 LOGISTIC REGRESSION**

This type of statistical analysis (also known as logit model) is often used for predictive analytics and modeling, and extends to applications in machine learning. In this analytics approach, the dependent variable is finite or categorical: either A or B (binary regression) or a range of finite options A, B, C or D (multinomial regression). It is used in statistical software to understand the relationship between the dependent variable and one or more independent variables by estimating probabilities using a logistic regression equation.

This type of analysis can help you predict the likelihood of an event happening or a choice being made. For example, you may want to know the likelihood of a visitor choosing an offer made on your website — or not (dependent variable). Your analysis can look at known characteristics of visitors, such as sites they came from, repeat visits to your site, behavior on your site (independent variables). Logistic regression models help you determine a probability of what type of visitors are likely to accept the offer — or not. As a result, you can make better decisions about promoting your offer or make decisions about the offer itself.

**3.2.3 SUPPORT VECTOR MACHINE (SVM)**

Support Vector Machine is kind of classification technique. It is a model used for classification and regression problems. It can solve linear and non-linear problems. The idea of SVM is simple: The algorithm creates a line or a hyper plane which separates the data into classes. When unknown data is given as input it predicts which class it belongs to. The margin between the hyper plane and the support vectors are as large as possible to reduce the error in classification.

**3.2.3 KNN (K-NEAREST NEIGHBOURS)**

KNN regression is a non-parametric method that, in an intuitive manner, approximates the association between independent variables and the continuous outcome by averaging the observations in the same neighbourhood. The size of the neighbourhood needs to be set by the analyst or can be chosen using cross-validation (we will see this later) to select the size that minimises the mean-squared error. While the method is quite appealing, it quickly becomes impractical when the dimension increases, i.e., when there are many independent variables.

**3.2.4 DECISION TREE**

As the name implies all decision tree techniques recursively separate observations into branches to construct a tree for the purpose of improving the prediction accuracy. Decision

tree is a conventional algorithm used for performing classifications based on the decisions made in one stage. This provides tree structured representation of the decision sets

**3.2.5 Nureal Network**

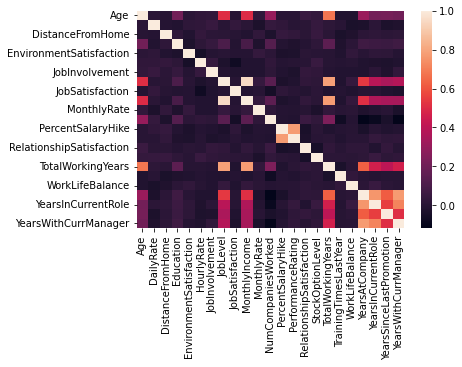
A neural network is a series of algorithms that endeavors to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates. In this sense, neural networks refer to systems of neurons, either organic or artificial in nature.

**3.2.6 Random Forest**

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of **ensemble learning,** which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model

**3.3 DATASET ANALYSIS**

# Data collection refers to the collection of relevant data from all available sources to perform analysis. The data used for this employee attrition analysis was obtained from Kaggle Website. This data set contains 1470 records and 35 attributes. The categorical values are converted to numeric values in order to make the classification algorithm more effectual. For example, categorical attribute ‘Business Travel’ contains three values such as Travel-Rarely, Travel-Frequently, Non-Travel. Hence it is converted to 1, 2 and 3 respectively. In this work, we used a publicly available dataset of HR details. This dataset is a simulated dataset that is created by IBM Watson Analytics (McKinley Stacker, 2015). This dataset contained standard HR features such as attrition, age, gender, education, last promotion, job title, and so on. This dataset con-tained 1470 employee records with 38 features. In this dataset 237 employee has “yes” attrition category while 1233 employee was “no” attrition category. Here all non-numeric values were assigned numerical values. The data conversion was performed using label encoding via the Scikit-learn package in Python (Pedregosa et al., 2011). Furthermore, Python is used in this work to train and evaluate various machine learning approaches. The correlation of different features is a heatmap. It shows how different features of the HR dataset are correlated. It also shows poorly correlated features and highly correlated features.



The Heatmap which helps to identify attributes with respective correlation.

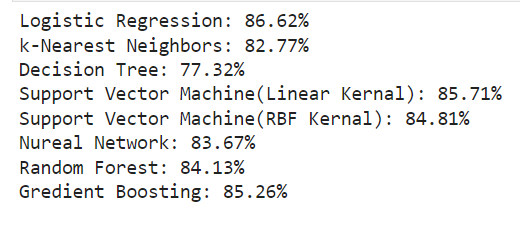
**3.4 RESULT AND EVALUATION**

In above dataset, there are various attributes like department, gender, overtime, business travel, etc. Based on these values, model which was build with the help of different machine learning algorithms which will predict whether employees will leave the organization or not. The predicted values are compared with test values to calculate the accuracy of the each algorithm. The table given below describes various factors, so we can easily

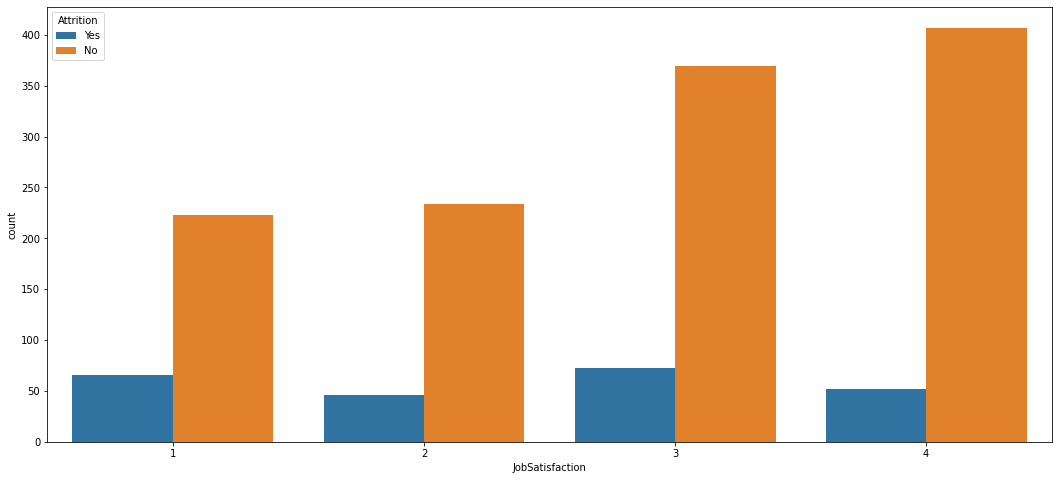
conclude which algorithm is best for our model. From the table, we can infer that Random Forest gives highest accuracy on the HR-Employee-Attrition dataset whereas Decision Tree gives the lowest accuracy for the same dataset. This paper presented the effect of voluntary attrition on organizations, and why predicting it is important. It further outlined various classification algorithms based on supervised learning to solve the prediction problem. The results of this research showed the superiority of the KNN classifier in terms of accuracy and predictive effectiveness, by means of the ROC curve. When used with its optimal configuration, it is a robust method that delivers accurate results in spite of the noise in the dataset, which is a major challenge for machine learning algorithms. The authors thus

recommend the use of the KNN classifier for accurately predicting employee attrition in an organization, which enables HR to take necessary action for the retention of employees

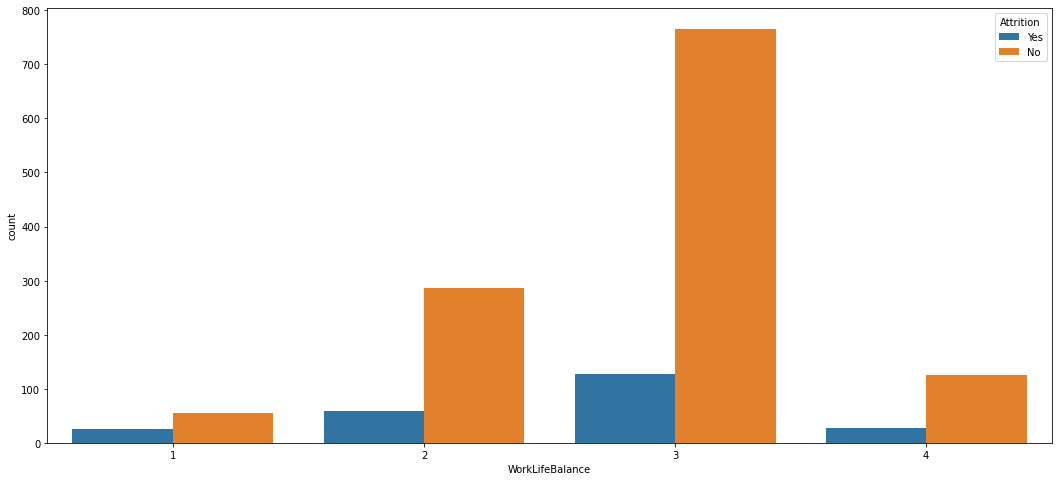
predicted to be at risk of leaving.



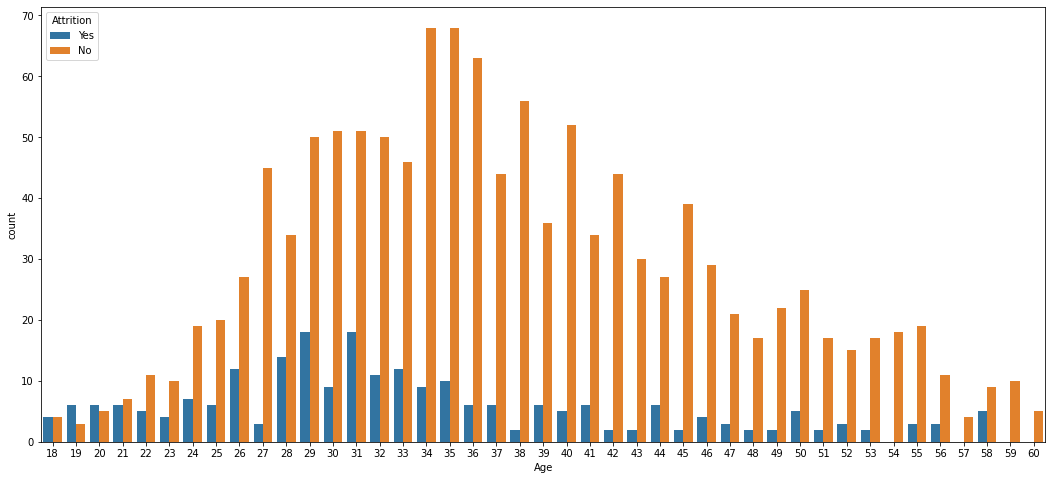
Since, Logistic Regression has the highest accuracy.



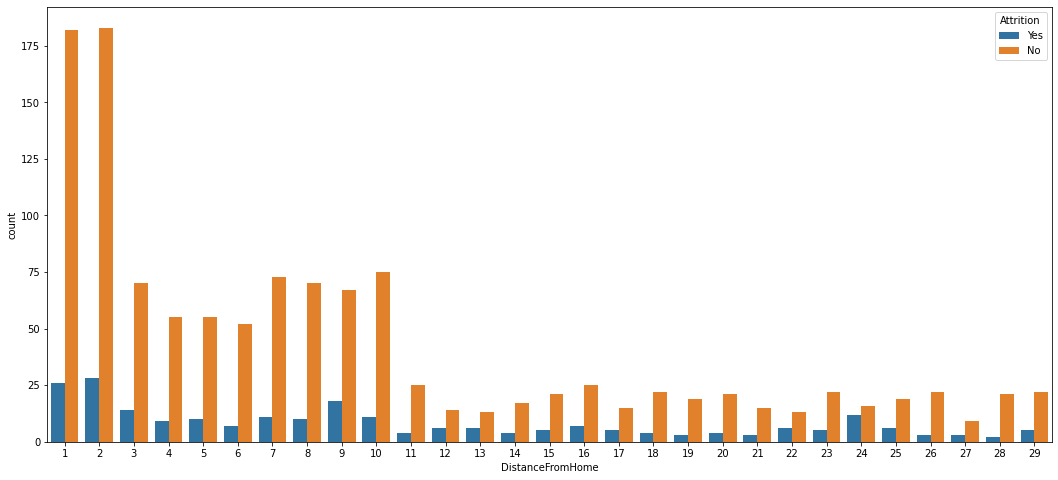
Attrition (vs) Job Satisfaction



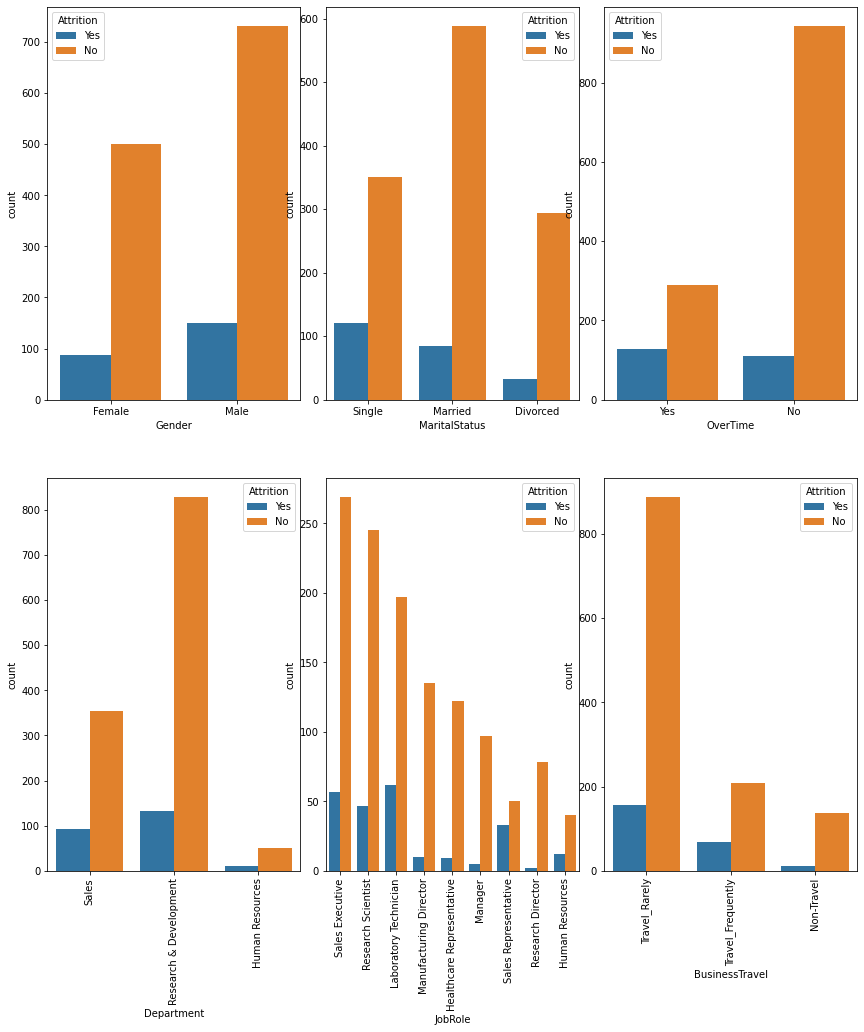
Attrition (vs) WorkLifeBalance



Attrition (vs) Age



Attrition (vs) Distance From Home

****

Attrition (vs) Categorical data

**CHAPTER 4**

**DESIGN**

**UML diagrams**

The Unified Modeling Language allows the software engineer to express an analysis model using the modeling notation that is governed by a set of syntactic semantic and pragmatic rules.

A UML system is represented using five different views that describe the system from distinctly different perspective. Each view is defined by a set of diagram, which is as follows.

* **User Model View**

1)This view represents the system from the users perspective.

2)The analysis representation describes a usage scenario from the end-users perspective.

* **Structural Model view**

1)In this model the data and functionality are arrived from inside the system.

2)This model view models the static structures.

* **Behavioral Model View**

It represents the dynamic of behavioral as parts of the system, depicting the interactions of collection between various structural elements described in the user model and structural model view.

* **Implementation Model View**

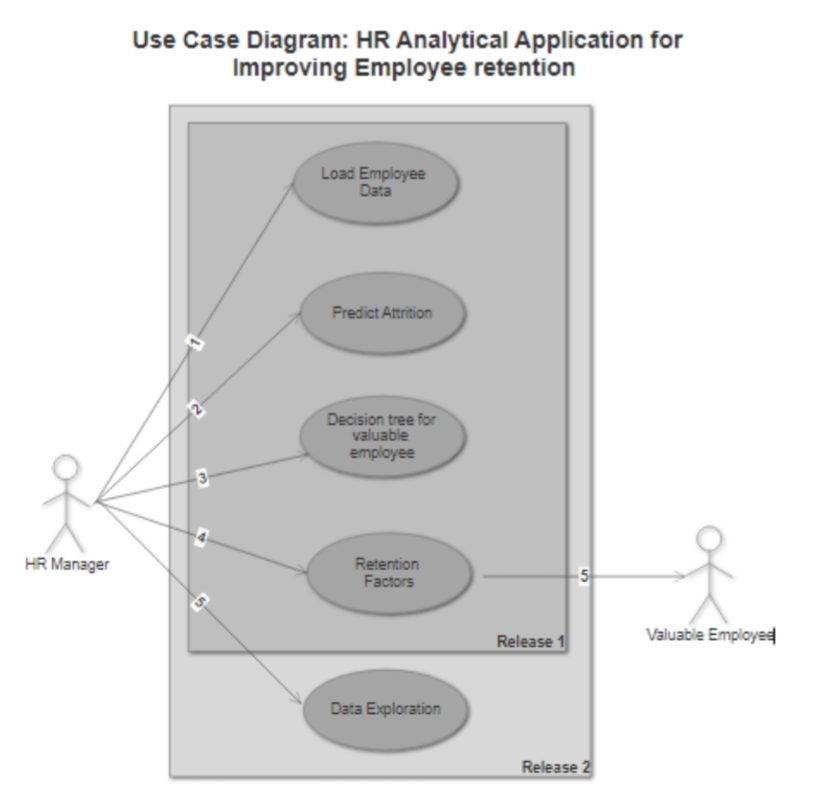
In this the structural and behavioral as parts of the system are represented as they are to be built.

* **Environmental Model View**

In this the structural and behavioral aspects of the environment in which the system is to be implemented are represented.

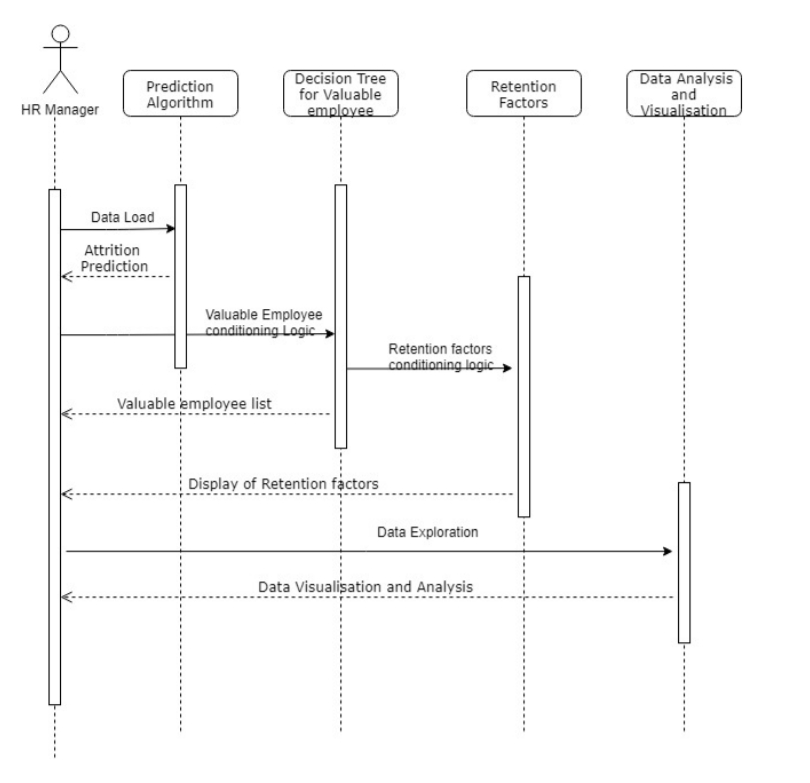
**4.1 USECASE DIAGRAM**

A **use case diagram** at its simplest is a representation of a user's interaction with the system and depicting the specifications of a use case. A use case diagram can portray the different types of users of a system and the various ways that they interact with the system. This type of diagram is typically used in conjunction with the textual use case and will often be accompanied by other types of diagrams as well.



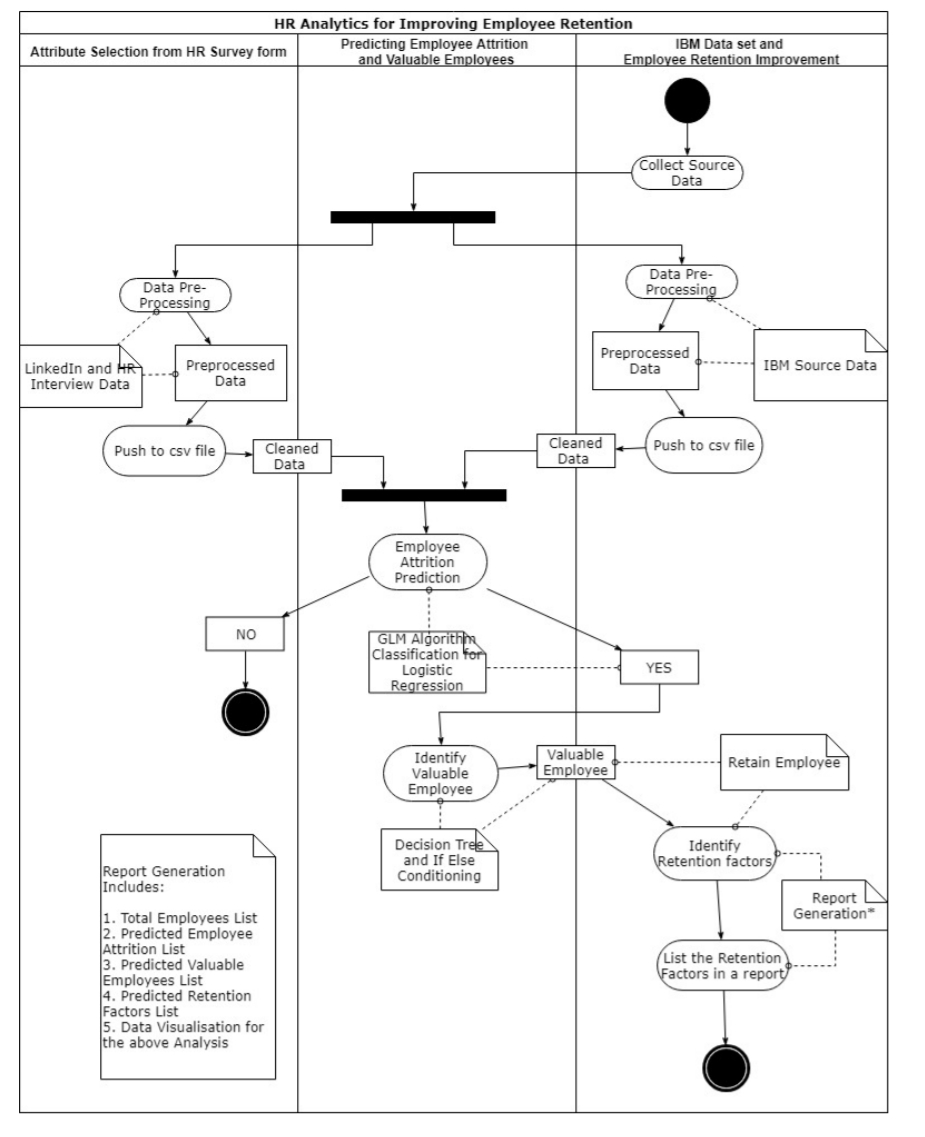
**4.2 Sequence Diagram**

A **sequence diagram** is a kind of interaction diagram that shows how processes operate with one another and in what order. It is a construct of a Message Sequence Chart. A sequence diagram shows object interactions arranged in time sequence. It depicts the objects and classes involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of the scenario. Sequence diagrams are typically associated with use case realizations in the Logical View of the system under development. Sequence diagrams are sometimes called **event diagrams**, **event scenarios**, and timing diagrams.



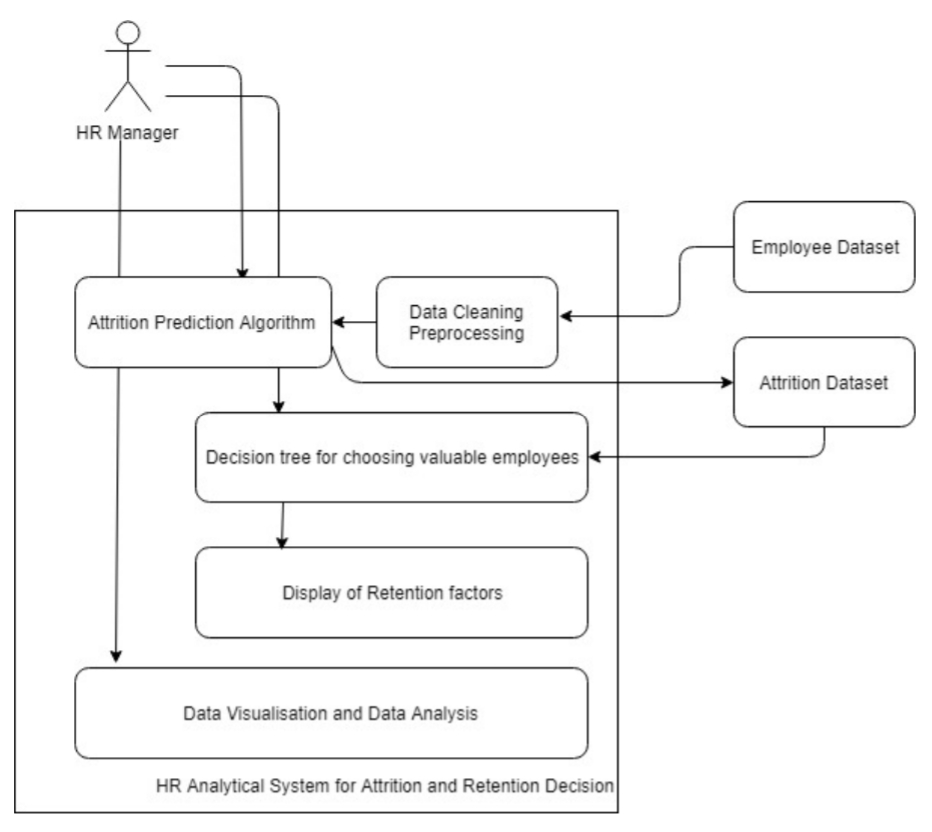
**4.3 Activity Diagram**

Activity diagram is another important diagram in UML to describe dynamic aspects of the system. It is basically a flow chart to represent the flow form one activity to another activity. The activity can be described as an operation of the system. So the control flow is drawn from one operation to another. This flow can be sequential, branched or concurrent.



**4.4 System Diagram**

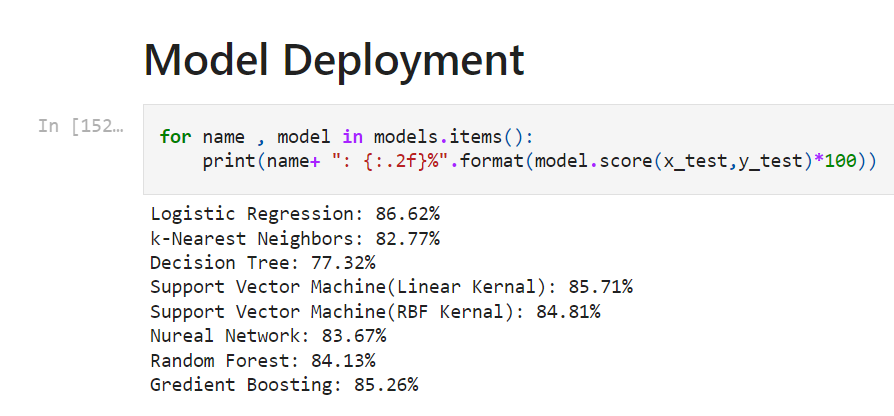
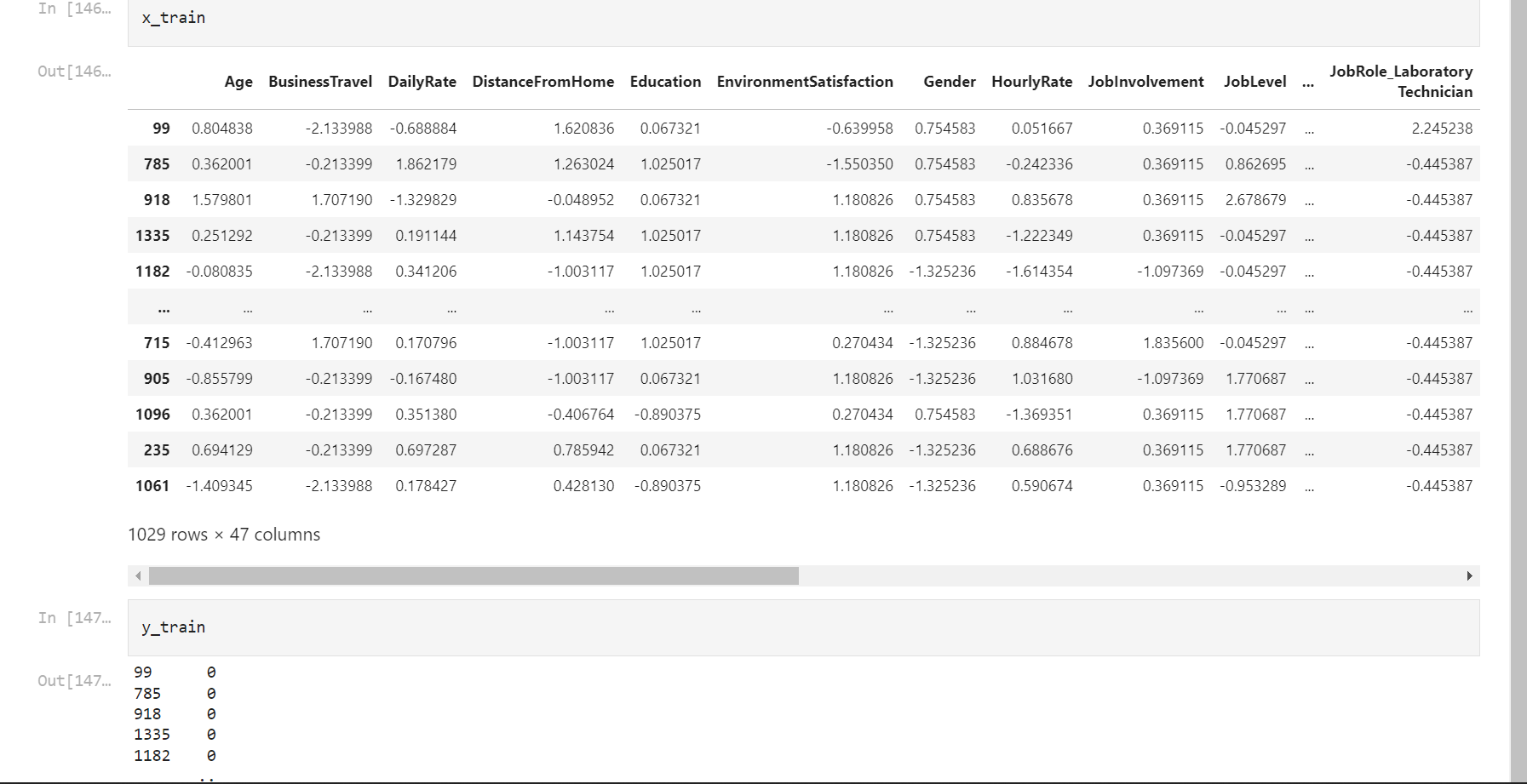
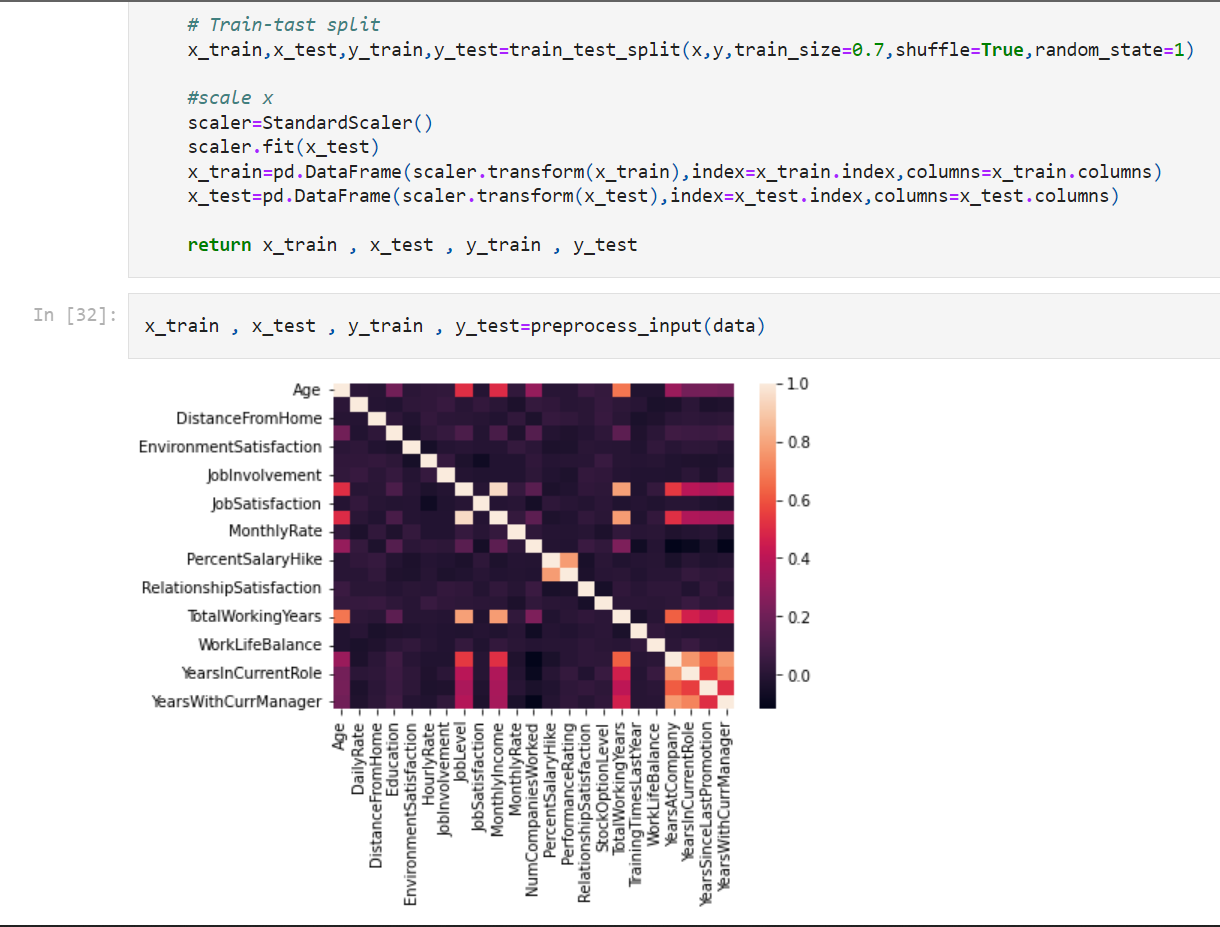
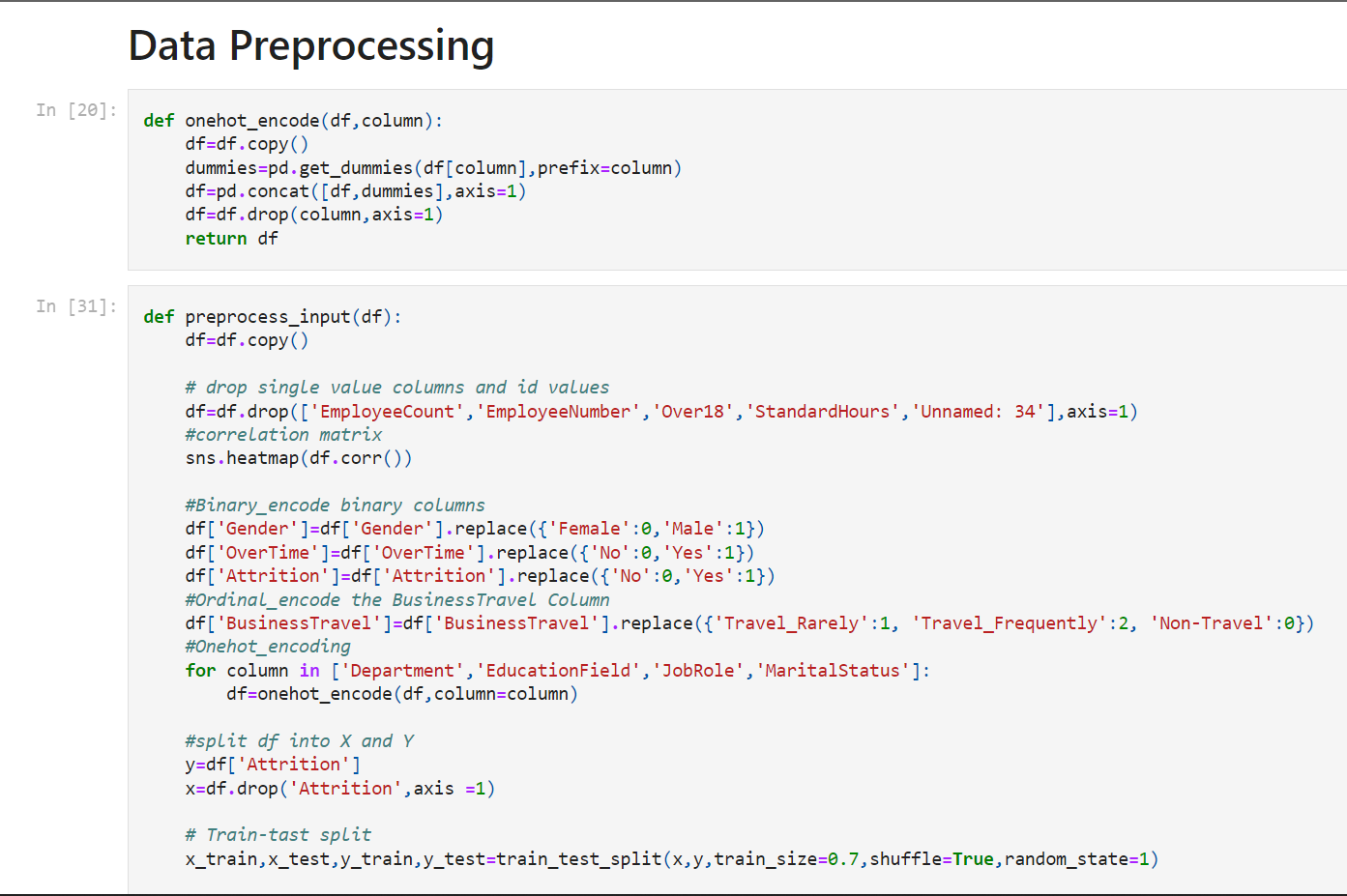
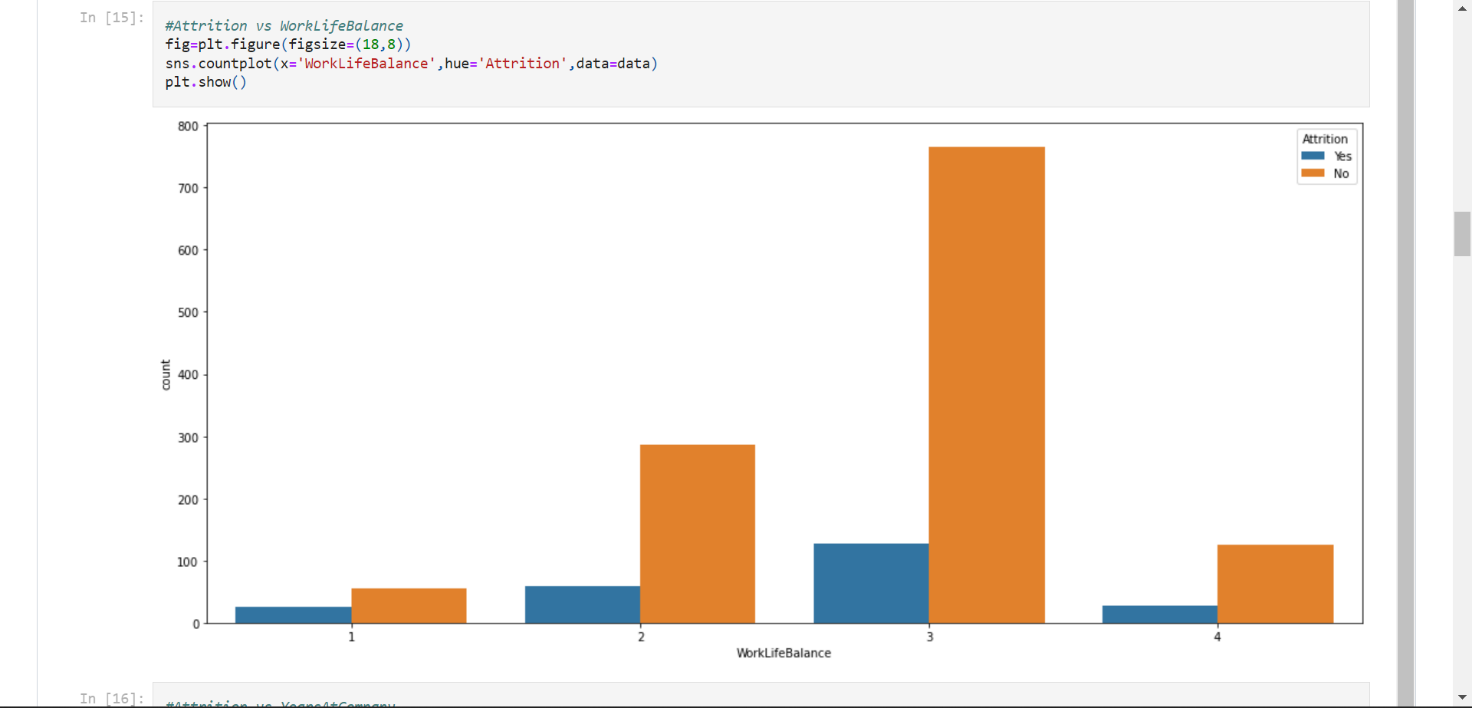
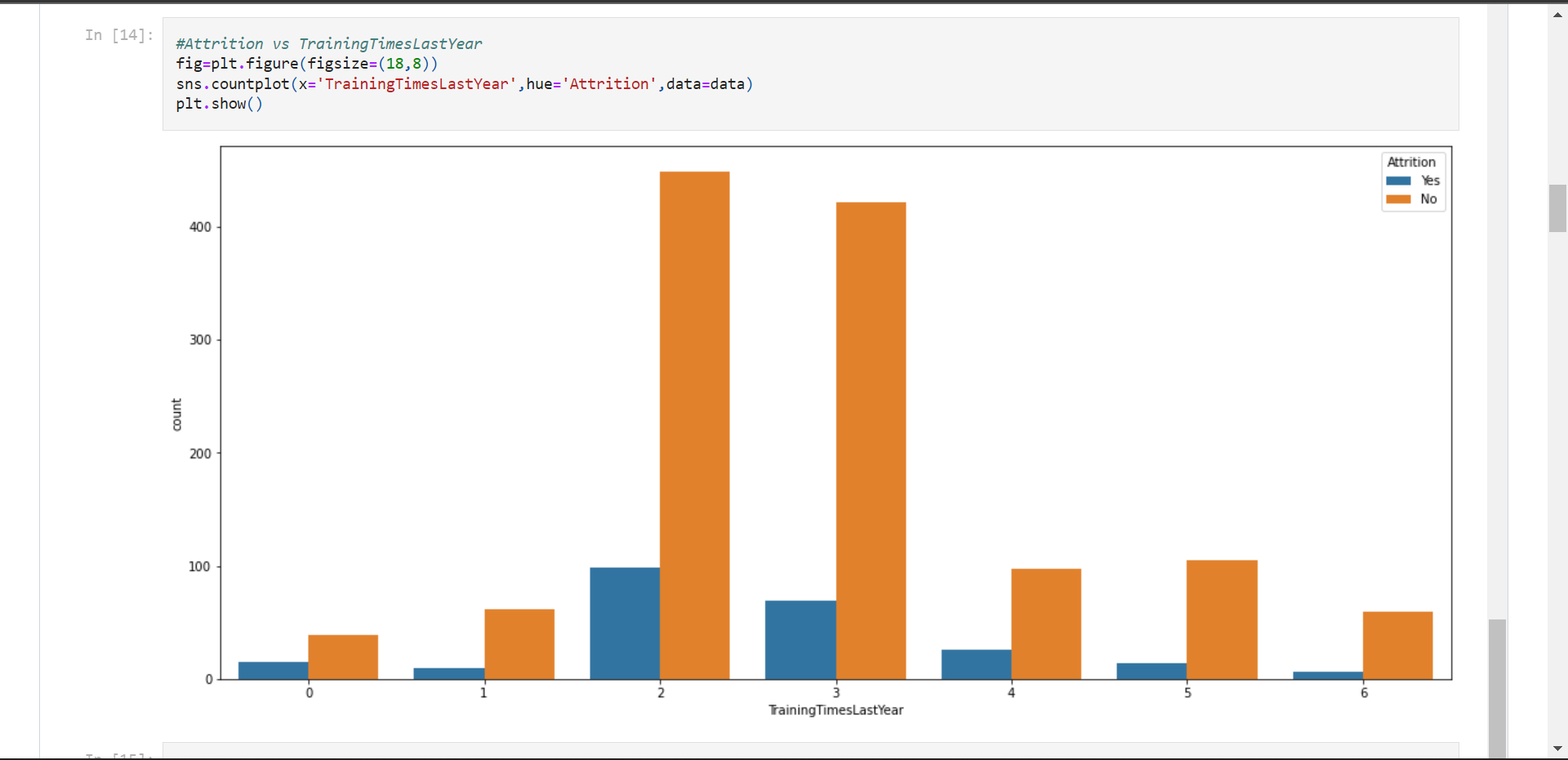
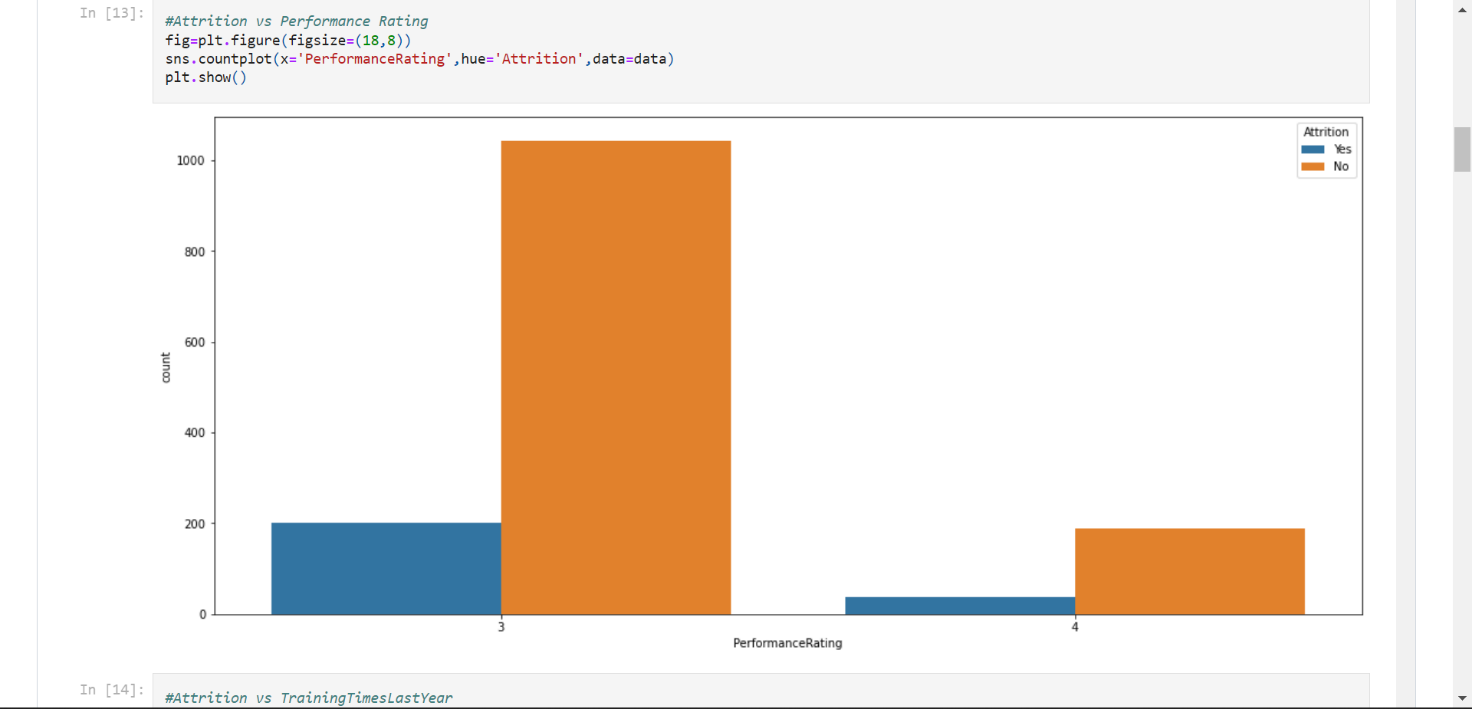
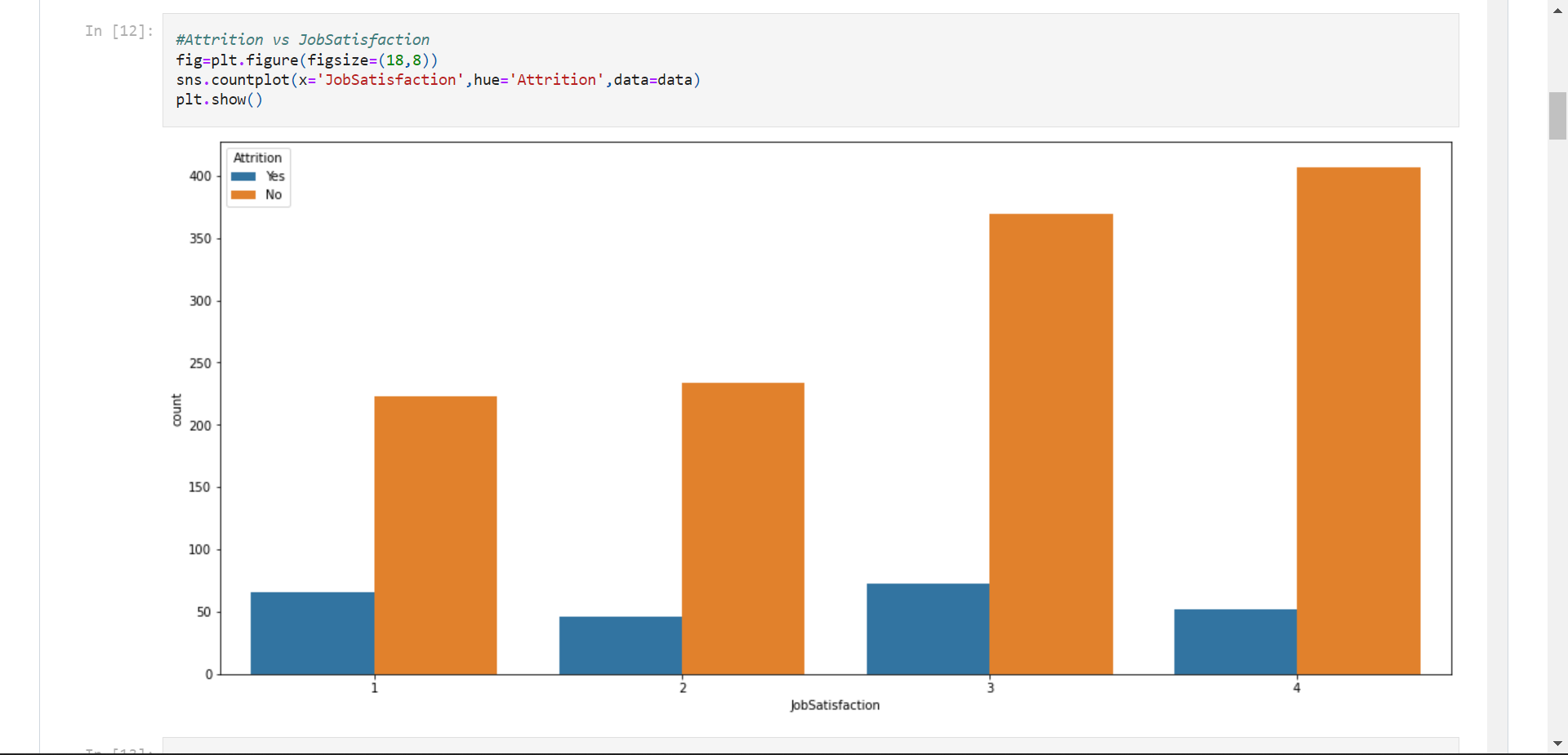
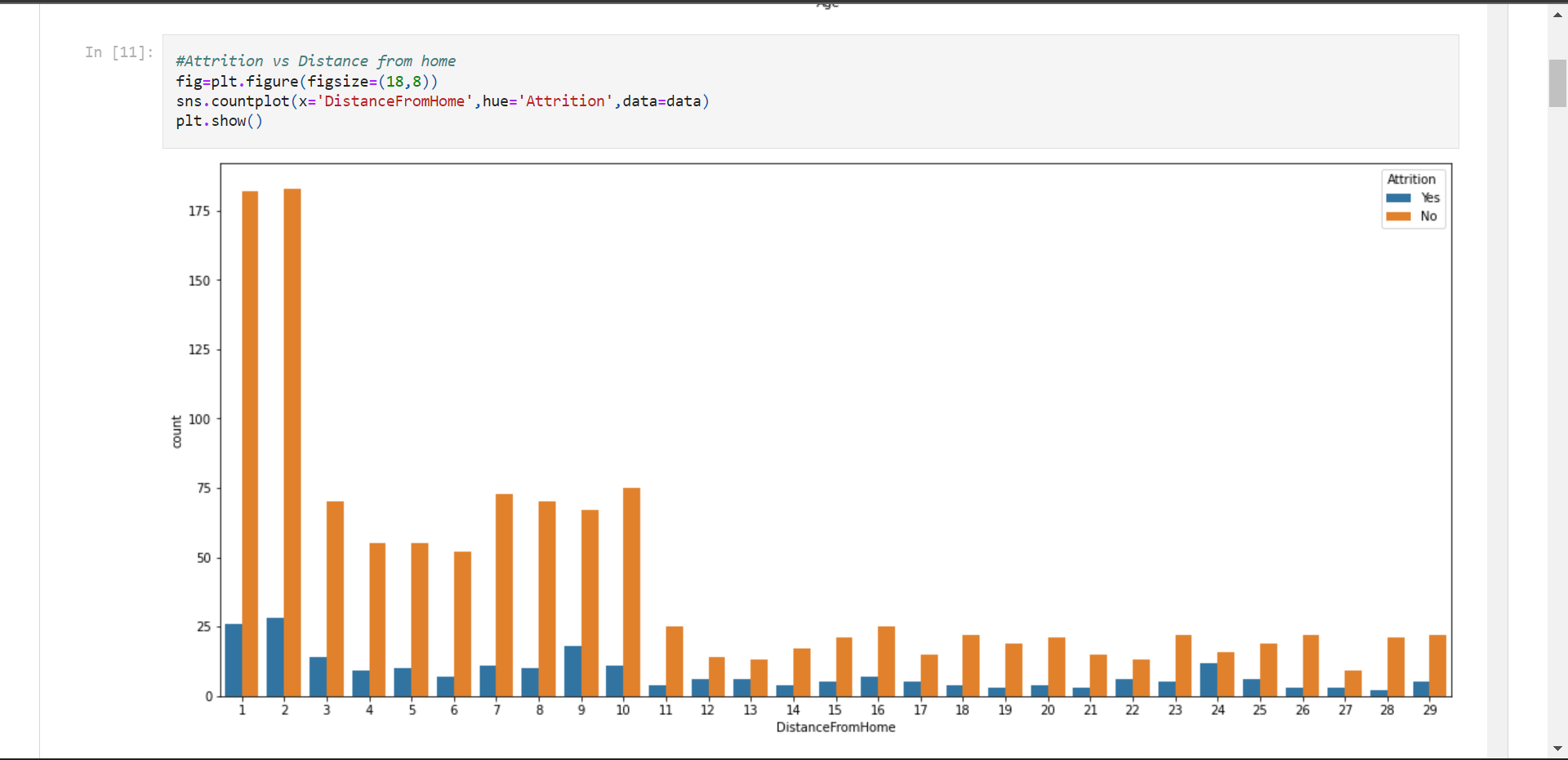
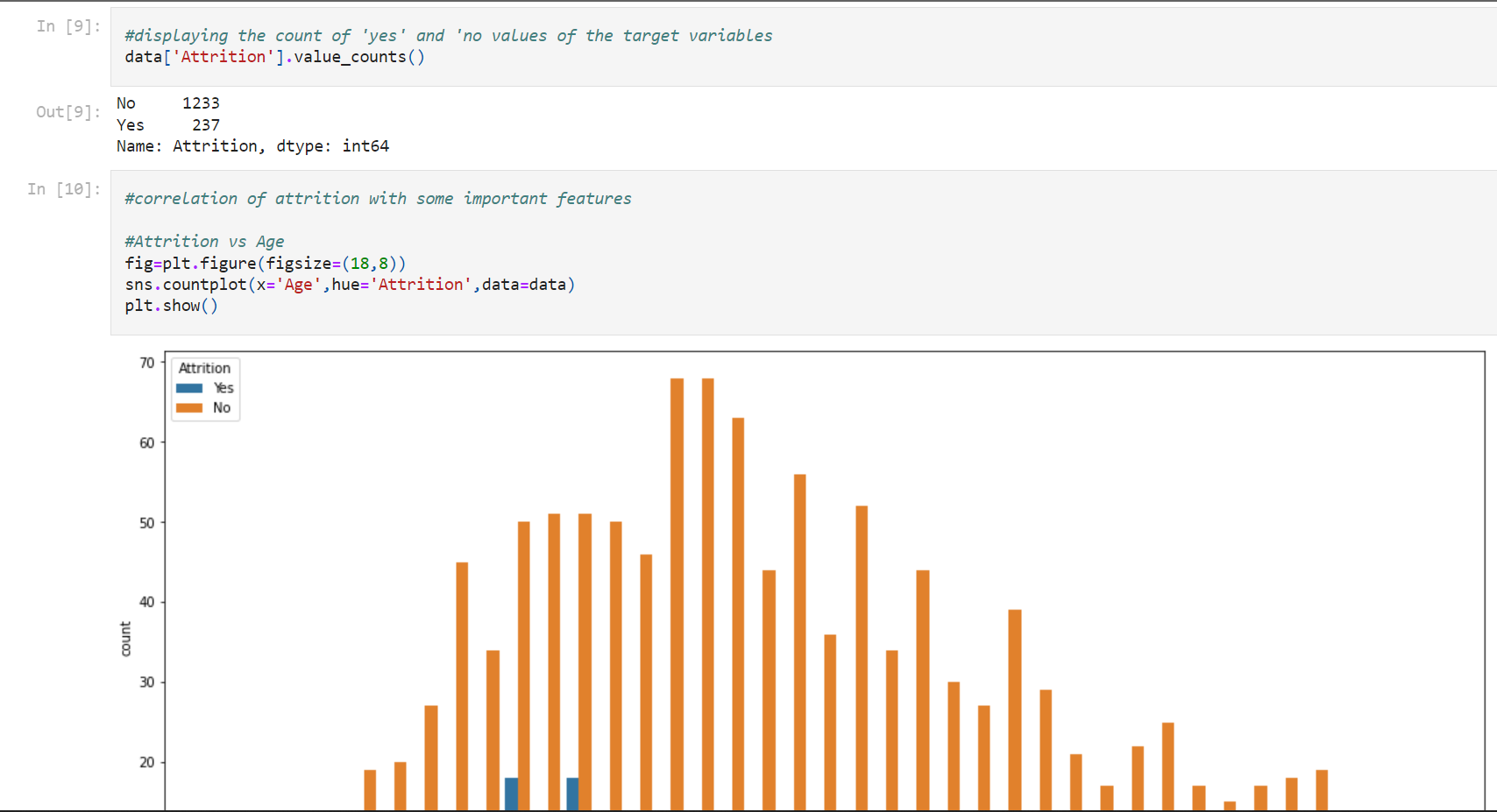
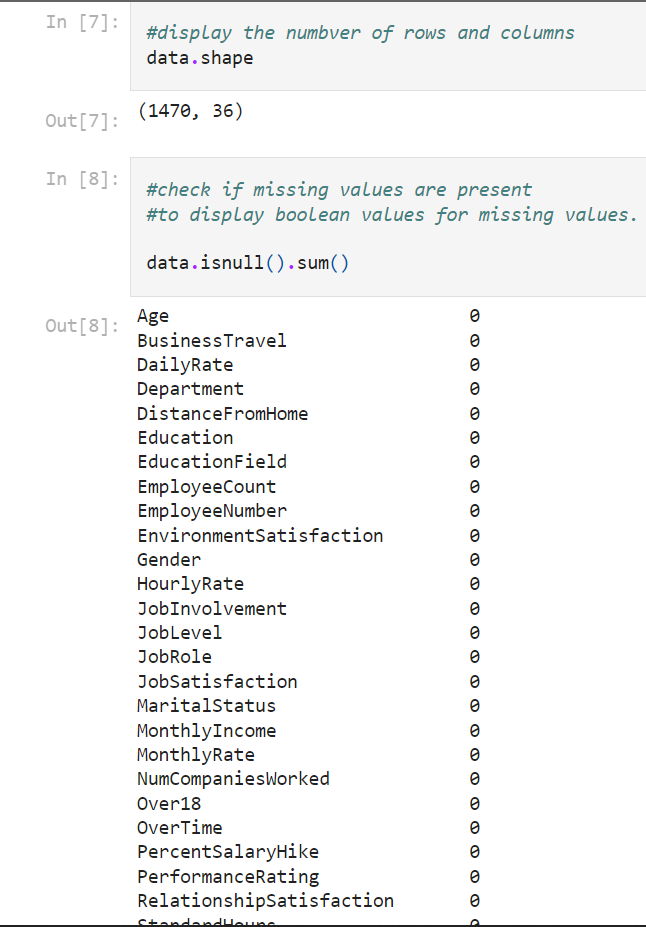
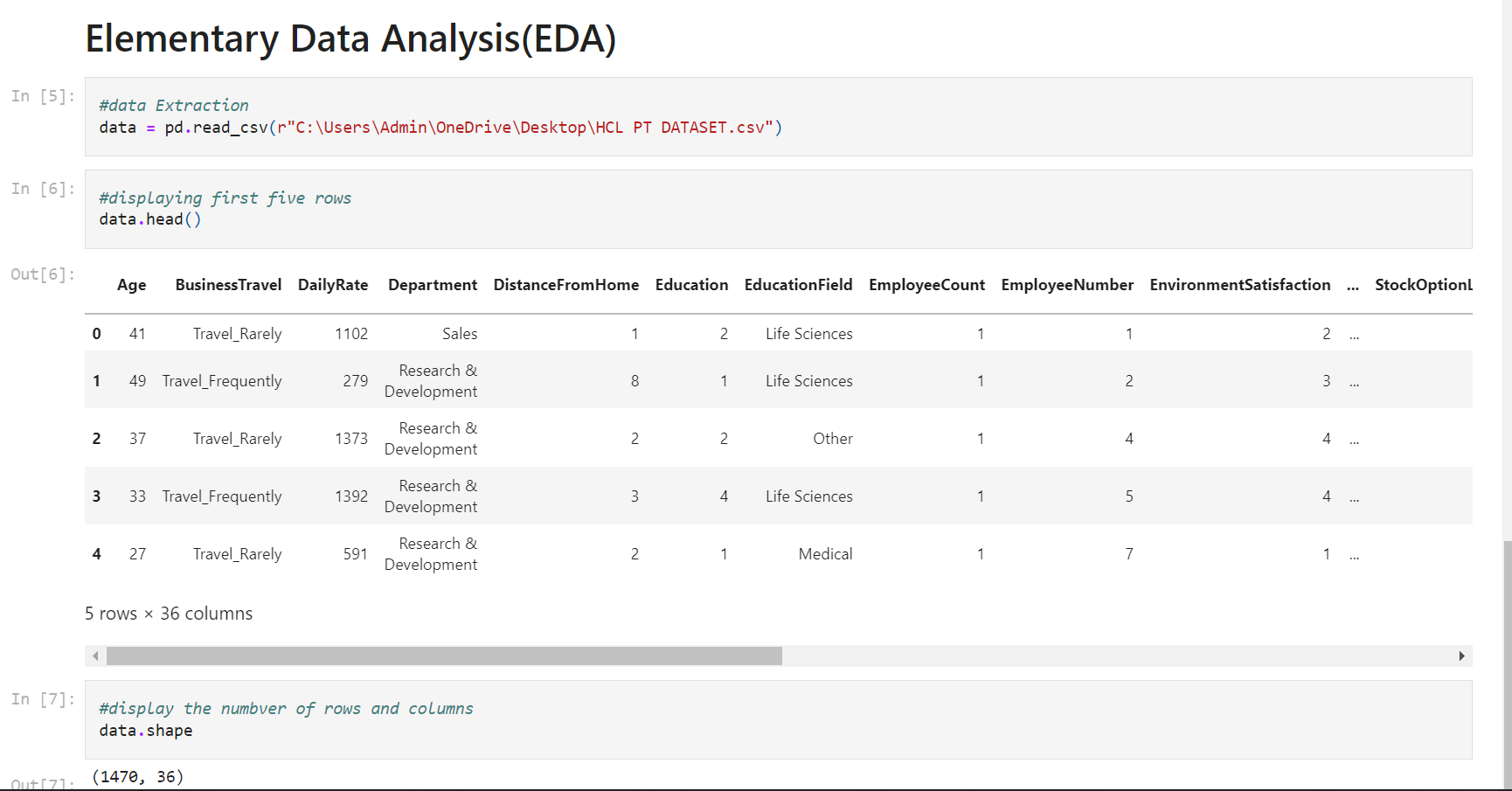
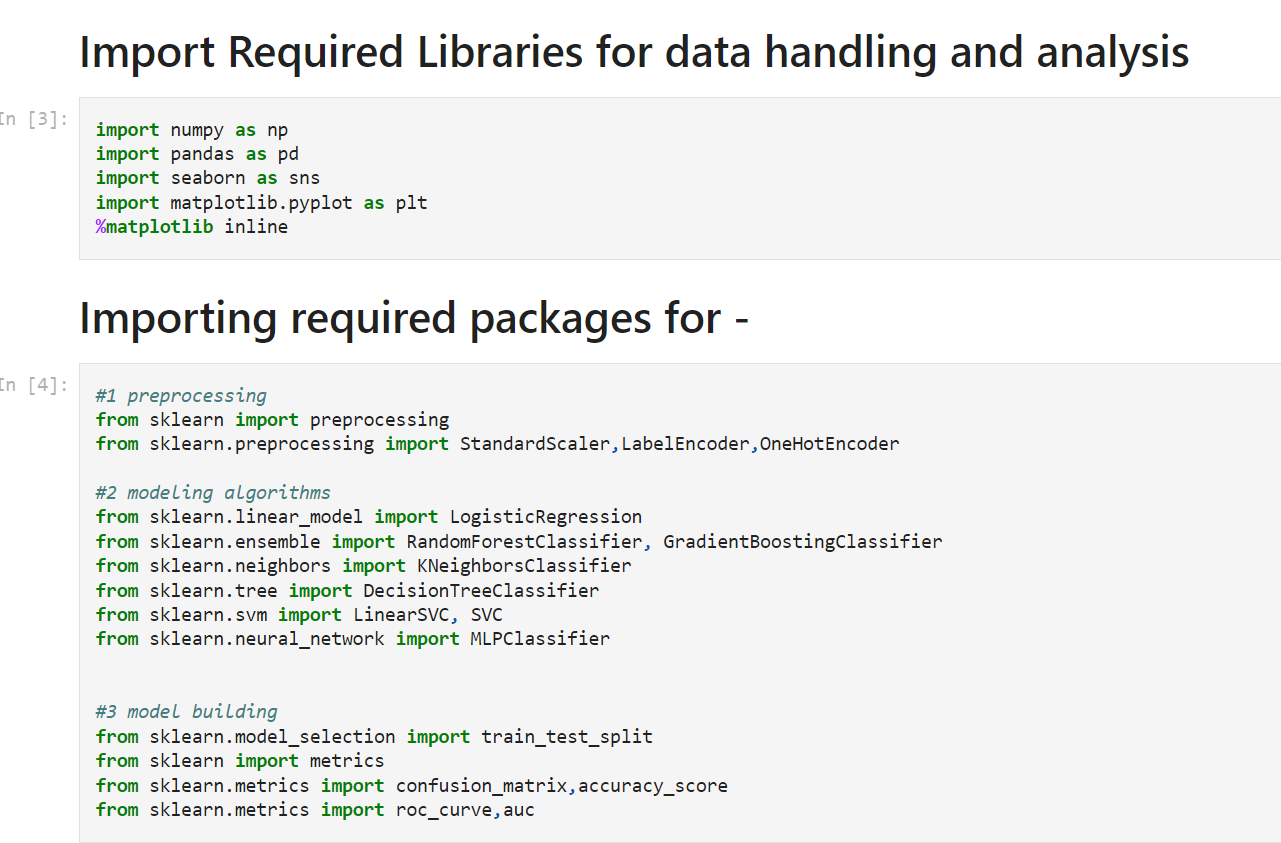
The system diagram below, shows the system design of an application which illustrates how the raw data is provided to the system and then retrieved as an output and again fed to system as a second input and get the final report (dashboard result in this case).



**CHAPTER 5**

**SCREEN SHOTS**

**5.1 ML CODE SCREEN SHOTS**



**CONCLUSION**

This chapter concludes the research work and recommends on the future advancements of the application and research. Thus, it can be seen from the above Literature reviews how the study on previous research work and case studies helped to find the attributes for predicting attrition, know the valuable employee and select the retention factors. From the study it is found that the factors responsible for attrition and retention are - Attrition, Percent Salary Hike, Monthly Income, Years Since Last Promotion, Distance From Home, Job Role, Performance Rating, Job Level, Environment Satisfaction, Years In Current Role, Relationship Satisfaction, Years With Current Manager, Job Satisfaction, Work Life Balance, Number of Companies Worked, Years At Company, Over-Time, Total Working Years, Marital Status, Age and Gender. Methodology demonstrated the processes followed while carrying on the research and developing an analytical application prototype. Finally, an Artifact design illustrated the development and working of system and application architecture. The advancement for improving the result accuracy is also mentioned so that it can be used for developing the application for commercial purpose. In this research, it is found that with the above attributes and Logistic regression algorithm, most accurate prediction result is obtained if the training dataset is 80% of the total data.

As a recommendation, the above analytical application can be integrated with the Human resource management finance budgeting application and there by predict the overall profit or savings in Human Resource Management process which include attrition, retention, hiring of new employee, amount spent on the training and development of new employee and loss to the project due to loss of valuable employee and prescribe on the further actions to be taken Thus, company can keep track of the amount of budget it had spent on Human resource management and budget to be spent on future and take necessary actions. Also, the system can be put on cloud and the data can directly be taken from the cloud storage through server connections by using FTP and SFTP commands in Unix environment.

**REFERENCES**

* **Ananya Sarker, S. S. D. M. S. Z. M. M. R., 2018. Employee’s Performance Analysis and Prediction using K-Means Clustering & Decision Tree Algorithm. Global Journal of Computer Science and Technology, 18(1), pp. 1-5.**
* **Archita Banerjee, R. K. G. M. G., 2017. A Study on the Factors Influencing the Rate of Attrition in IT Sector: Based on Indian Scenario. Pacific Business Review International, 9(7), p. 01.**
* **Qiao, Zihan. (2020).Attrition analysis and prediction on machine learining algorithms. 76-79. 10.1109/MLBDBI51377.2020.00020**
* **Xie, Han & Li, Chaoyi & Ding, Ning & Gong, Changrui. (2021). Attrition analysis Based On Machine Learning model. 366-369. 10.1109/ICEKIM52309.2021.00087**
* **Saraswathi, K. & Renukadevi, N. & Nandhinidevi, S. & Gayathridevi, S. & Naveen, P.. (2021). Attrition prediction using machine learning approaches. AIP Conference Proceedings. 2387. 140038. 10.1063/5.0068655.**

**Website links**

<https://www.kaggle.com/code/vincentlugat/ibm-attrition-analysis-and-prediction>

**GITHUB LINK**

<https://github.com/Peravalideepesh/attrition-analysis-and-prediction>

Peravali Deepesh

39110770

SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY