

|  |
| --- |
| **INTRODUCTION TO ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING** |
| Lab Workbook |
| **Department of Computer Science and Engineering**  **The NorthCap University, Gurugram** |

**INTRODUCTION TO ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING**

**(CSL 236)**

**Lab Workbook**



**Faculty name**: Dr. Ruchika Lalit **Student name**: Mohit Khatana

**Roll No**.: 20CSU294

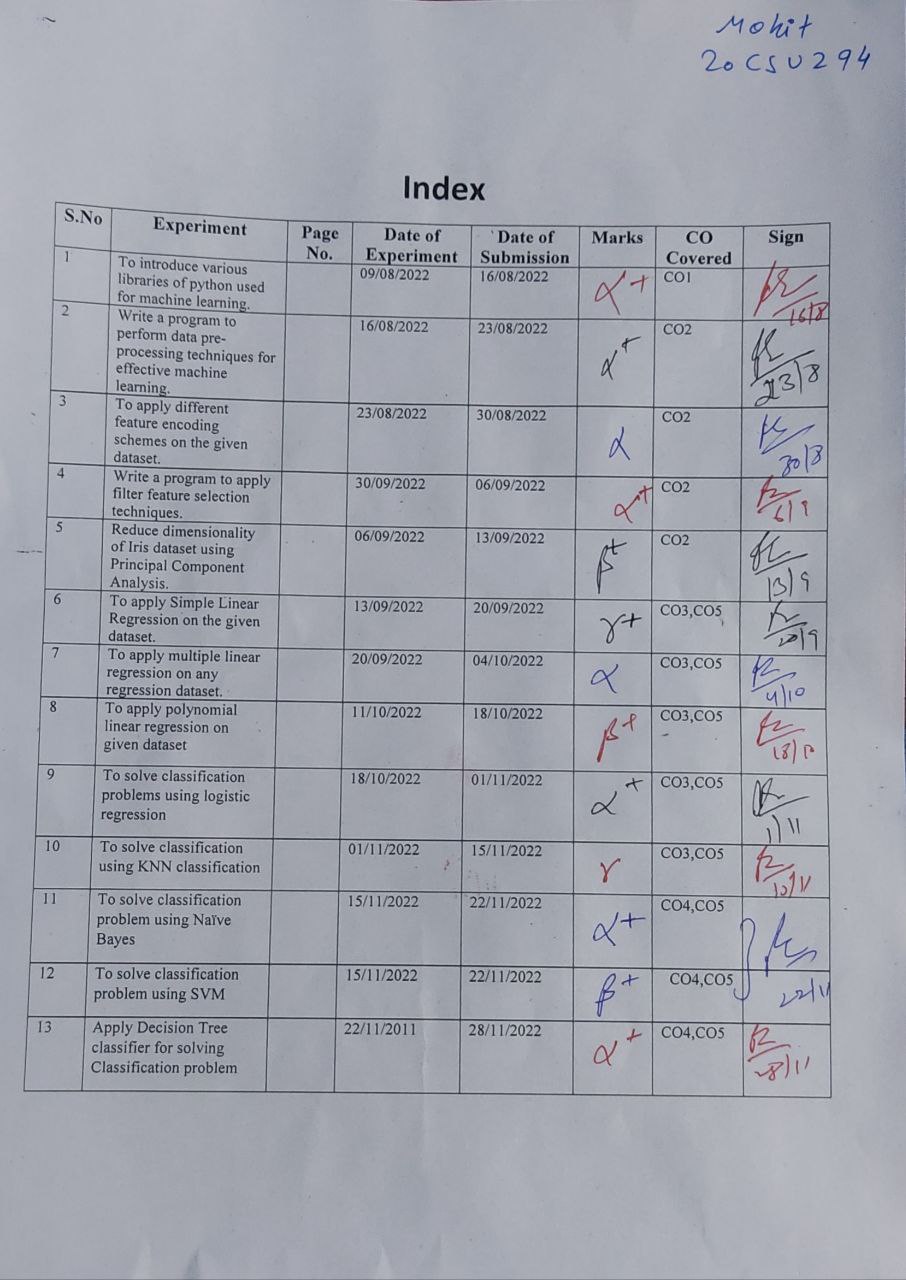
**Semester**: 5th

**Group**: AIML – B2

**Department of Computer Science and Engineering**

**NorthCap University, Gurugram- 122001, India**

**Session 2022-23**

****

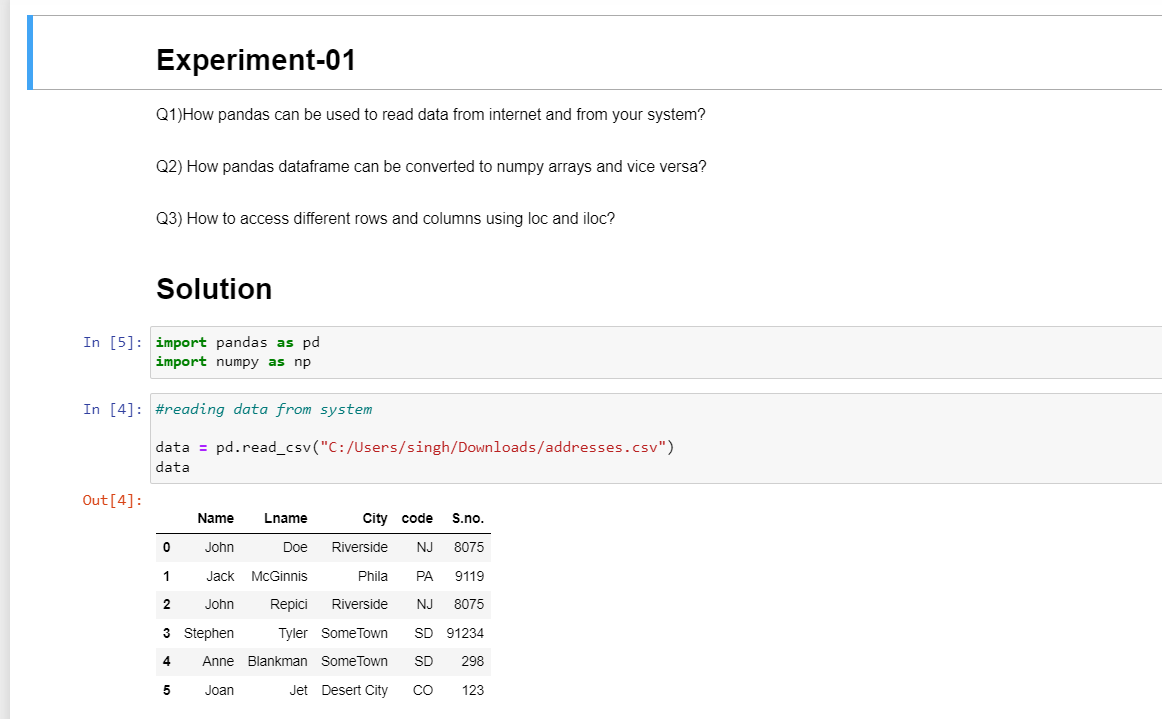
**EXPERIMENT NO. 1**

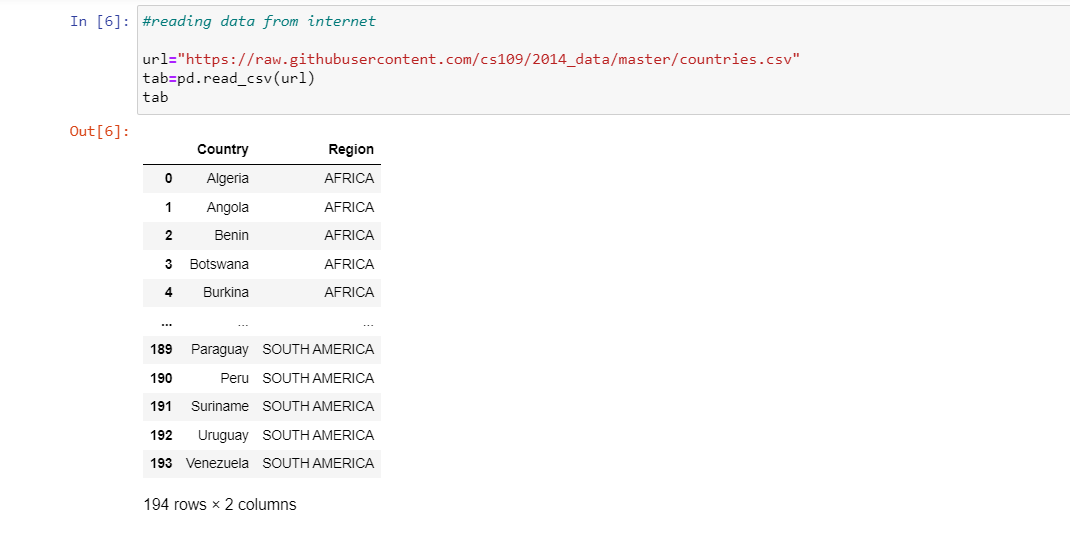
|  |
| --- |
| **Student Name and Roll Number:**  Mohit Khatana 20CSU294 |
| **Semester /Section:** 5 AIML B1 |
| **Link to Code:** |
| **Date: 04/08/22** |
| **Faculty Signature:** |
| **Grade:** |

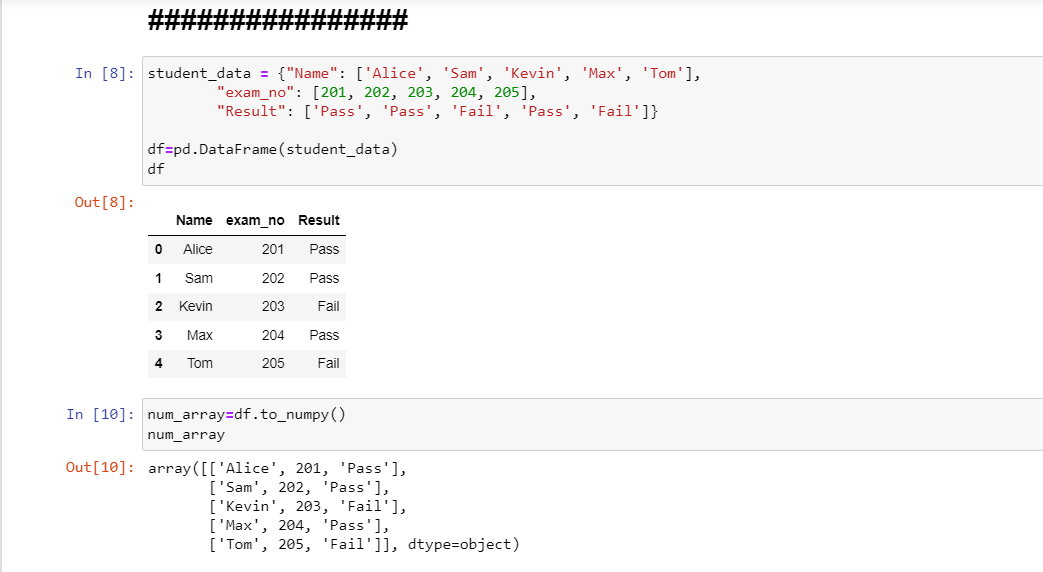
|  |
| --- |
| **Objective(s):**   * To understand the basic libraries of python. * To differentiate between numpys and pandas. |
| **Outcome:**  Students will be familiarized with handling dataset using python basic libraries and applying various operations on dataset using these. |
| **Problem Statement:**  To introduce various libraries of python used for machine learning. |
| **Background Study:**  **Basic libraries** of python are necessary to import datasets and applying various data pre-processing and machine learning techniques on them. |
| **Question Bank:**  1. How pandas can be used to read data from internet and from your system?  **Ans: pd.read\_csv(‘filepath’)**  2. How pandas dataframe can be converted to numpy arrays and vice versa?  **Ans. .toNumpy, and pd.DataFrame()**  3. How to access different rows and columns using loc and iloc?  **Ans. Loc can access columns by name and conditions, whereas iloc can access columns by column numbers and integer based conditions.**  4. Differentiate between Feature Selection and Dimensionality Reduction.  **Ans. Feature selection is simply selecting and excluding given features without changing them. Dimension reduction is the process of reducing the number of random variables under consideration.**  5. What are the advantages of Wrapper methods over filter methods for feature selection?  **Ans. Filter methods measure the relevance of features by their correlation with dependent variable while wrapper methods measure the usefulness of a subset of feature by actually training a model on it.**  6. Explain Regularization methods for Feature Selection.  **Ans. Regularisation consists in adding a penalty to the different parameters of the machine learning model to reduce the freedom of the model and in other words to avoid overfitting. In linear model regularisation, the penalty is applied over the coefficients that multiply each of the predictors.**  7. What are Embedded feature selection methods.  **Ans. Embedded methods combine the qualities of filter and wrapper methods. It's implemented by algorithms that have their own built-in feature selection methods. Some of the most popular examples of these methods are LASSO and RIDGE regression which have inbuilt penalization functions to reduce overfitting.** |

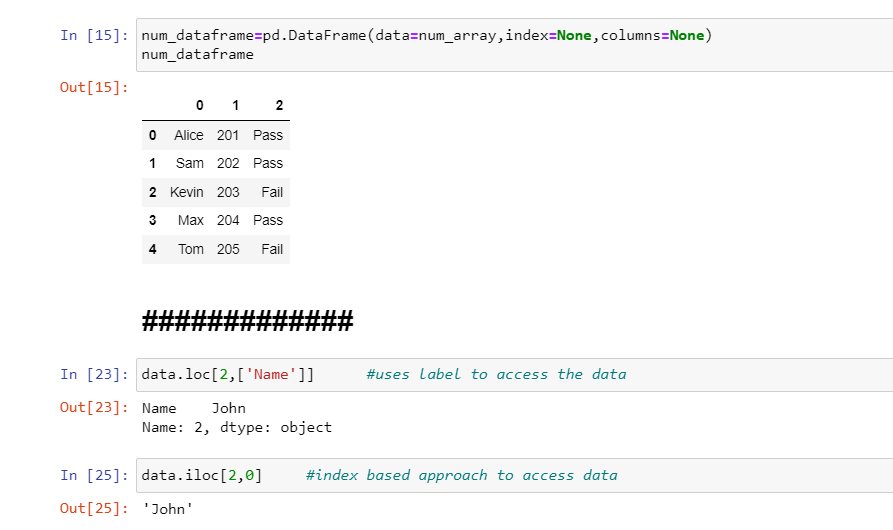
**Student Work Area**

**Algorithm/Flowchart/Code/Sample Outputs**









**EXPERIMENT NO. 2**

|  |
| --- |
| **Student Name and Roll Number:** Mohit Khatana 20CSU294 |
| **Semester /Section:**5TH| AIML - B |
| **Link to Code:** |
| **Date: 18/08/22** |
| **Faculty Signature:** |
| **Grade:** |

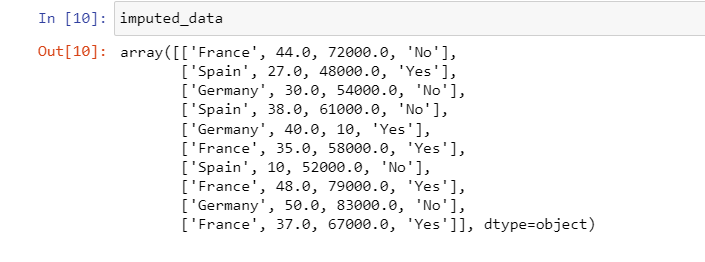
|  |
| --- |
| **Objective(s):**   * To understand the importance of data pre-processing techniques. * To handle missing values, duplicate values, feature scaling etc. |
| **Outcome:**  Students will be familiarized with the understanding and importance of applying various data pre-processing techniques. |
| **Problem Statement:**  Write a program to perform data pre-processing techniques for effective machine learning. |
| **Background Study:** Data preprocessing in Machine Learning is a crucial step thathelps enhance the quality of data to promote the extraction of meaningful insights from the data. Data preprocessing in Machine Learning refers to the technique of preparing (cleaning and organizing) the raw data to make it suitable for a building and training Machine Learning models. |
| **Question Bank:**  1. What are different ways to handle missing values both for numerical as well as categorical data?  **Ans)Numerical values: fillna, dropna, filling in the missing values with the average value of surrounding data.**  **Categorical values:**   1. **Ignore these observations.** 2. **Replace with general average.** 3. **Replace with similar type of averages.** 4. **Build model to predict missing values.**   2. What is the function in python used for finding duplicate rows in data?  **pandas.DataFrame.duplicated()**  3. Differentiate between two scaling methods used for feature scaling?  **1) Min Max Scaler: Standardizes the values according to min and max values of the dataset. 2) Standard Scaler: Standardize the values using z-score, also known as normalization. Ensures that the dataset’s mean is 0, and the std. deviation is 1.** |

**Student Work Area**

**Algorithm/Flowchart/Code/Sample Outputs**







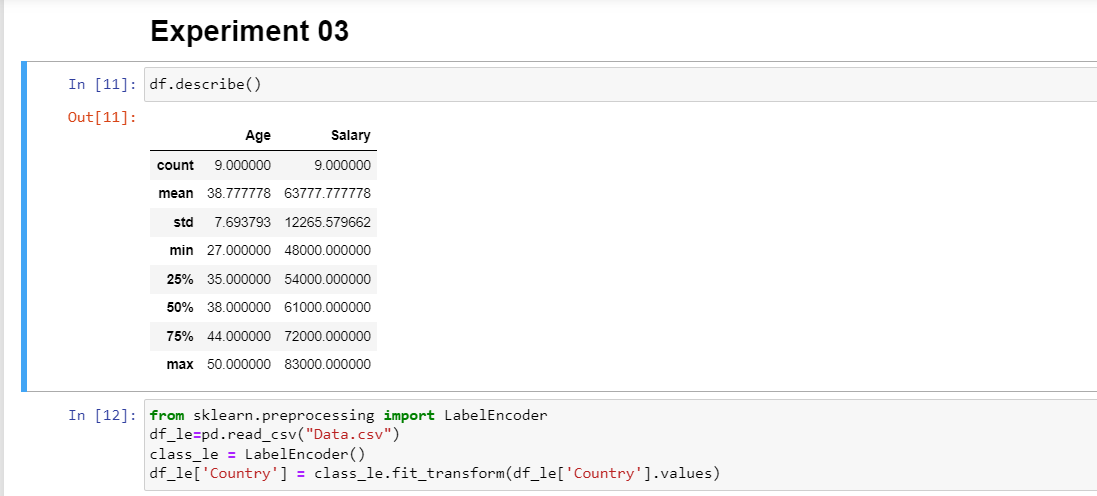
**EXPERIMENT NO. 3**

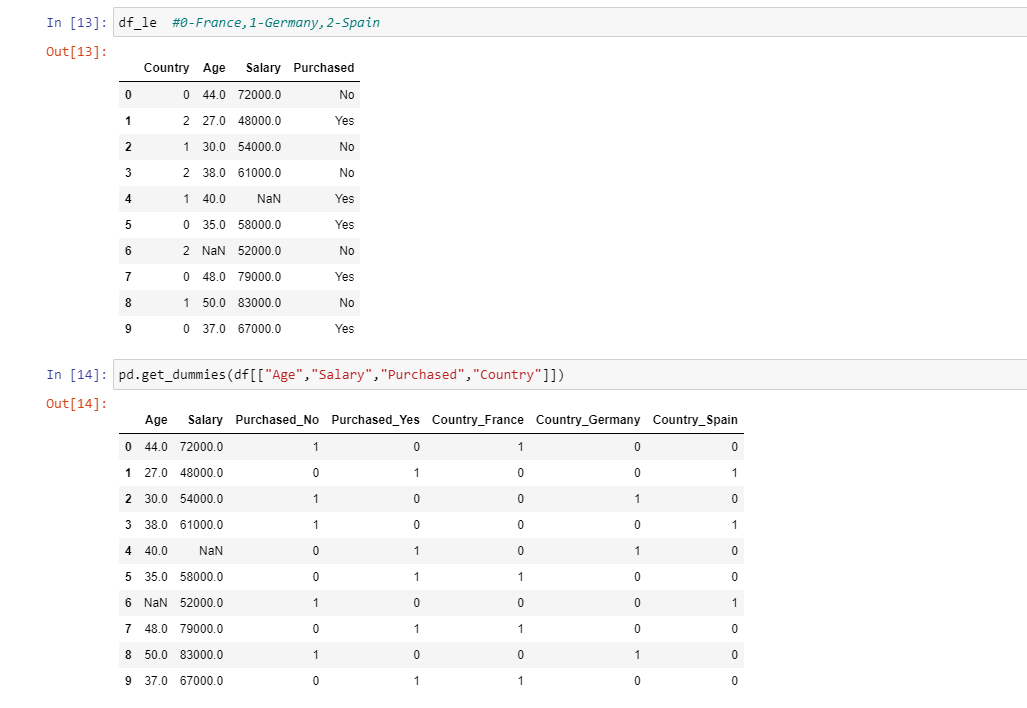
|  |
| --- |
| **Student Name and Roll Number:** Mohit Khatana 20CSU294 |
| **Semester /Section:** 5TH | AIML - B |
| **Link to Code:** |
| **Date: 26/08/22** |
| **Faculty Signature:** |
| **Grade:** |

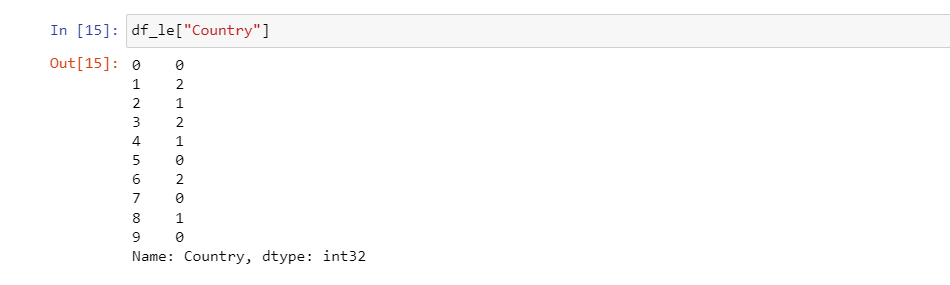
|  |
| --- |
| **Objective(s):**   * To perform different encoding schemes. * To prepare dataset by converting categorical data to numeric form for machine learning. * To understand and implement label encoding and one hot encoding. |
| **Outcome:**  Students will be able to understand different encoding schemes to prepare data for machine learning. |
| **Problem Statement:**  To apply different feature encoding schemes on the given dataset. |
| **Background Study:** Machine learning models require all input and output variables to be numeric. This means that if your data contains categorical data, you must encode it to numbers **before you can fit and evaluate a model**. The two most popular techniques are Label Encoding and One-Hot Encoding. |
| **Question Bank:**  1. Can ML algorithms handle categorical data directly?  **Ans)No, they required encoding techniques for converting the categorical data to numerical values.**  2. What are the different schemes for encoding categorical data?  **Some encoding schemes are:**   * **One-Hot Encoding** * **Binary Encoding** * **Frequency Encoding** * **Label encoding** * **Ordinal Encoding**   3. Differentiate between Label Encoding and One Hot Encoding?  **In label encoding, each label is assigned a unique integer based on alphabetical ordering.**  **In One-Hot encoding, each category is represented as one-hot vector.** |

**Student Work Area**

**Algorithm/Flowchart/Code/Sample Outputs**







**EXPERIMENT NO. 4**

|  |
| --- |
| **Student Name and Roll Number:** Mohit Khatana 20CSU294 |
| **Semester /Section5 AIML B1** |
| **Link to Code:** |
| **Date: 01/09/22** |
| **Faculty Signature:** |
| **Grade:** |

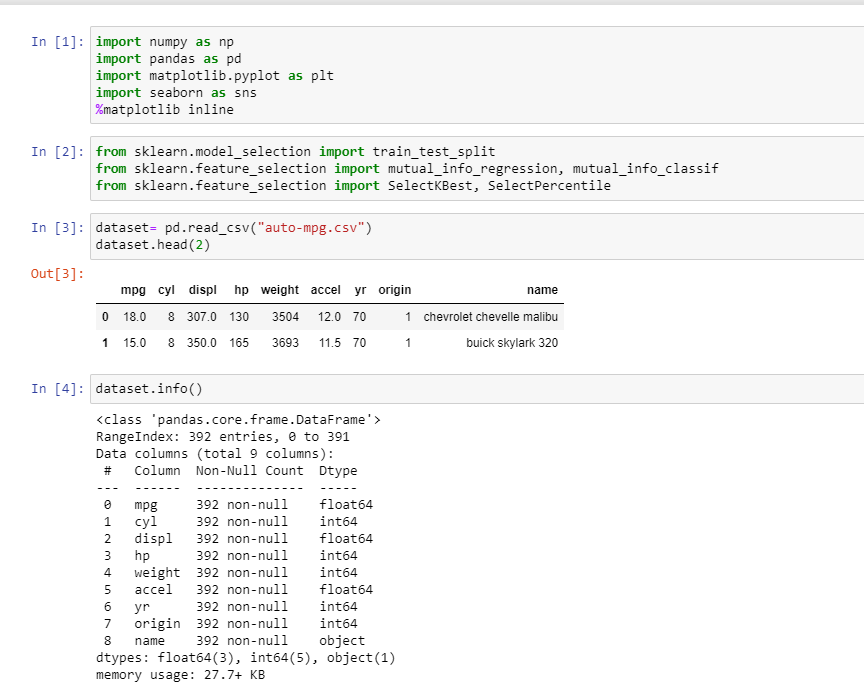
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| |  | | --- | | **Objective(s):**   * To understand the importance of feature selection * To differentiate between different types of feature selection. * Build a model using feature selection techniques. | | **Outcome:**  Students will be familiarized with model building using feature selection techniques and  optimization. | | **Problem Statement:**  Write a program to apply filter feature selection techniques. | | **Background Study:** Feature selection is the process of reducing the number of input variables when developing a predictive model. It is desirable to reduce the number of input variables to both reduce the computational cost of modeling and, in some cases, to improve the performance of the model. | | **Question Bank:**   1. What are different filter feature selection techniques?   Ans)   * **Information Gain** * **Chi-Square Test** * **Fisher’s score** * **Correlation Coefficient** * **Variance Threshold** * **Mean Absolute Difference** * **Dispersion Ratio**   2. How feature selection techniques depend on the data type of input features and output variable?  **Ans) Input variables are those that are provided as input to a model. In feature selection, it is this group of variables that we wish to reduce in size. Output variables are those for which a model is intended to predict, often called the response variable.**  **The type of response variable typically indicates the type of predictive modeling problem being performed. For example, a numerical output variable indicates a regression predictive modeling problem, and a categorical output variable indicates a classification predictive modeling problem.**   * **Numerical Output: Regression predictive modeling problem.** * **Categorical Output: Classification predictive modeling problem.**   3. What is the mathematics behind Pearson’s Correlation to rank features?  **Ans) The Pearson correlation for two objects, with paired attributes, sums the product of their differences from their object means and divides the sum by the product of the squared differences from the object means** | |

**Student Work Area**

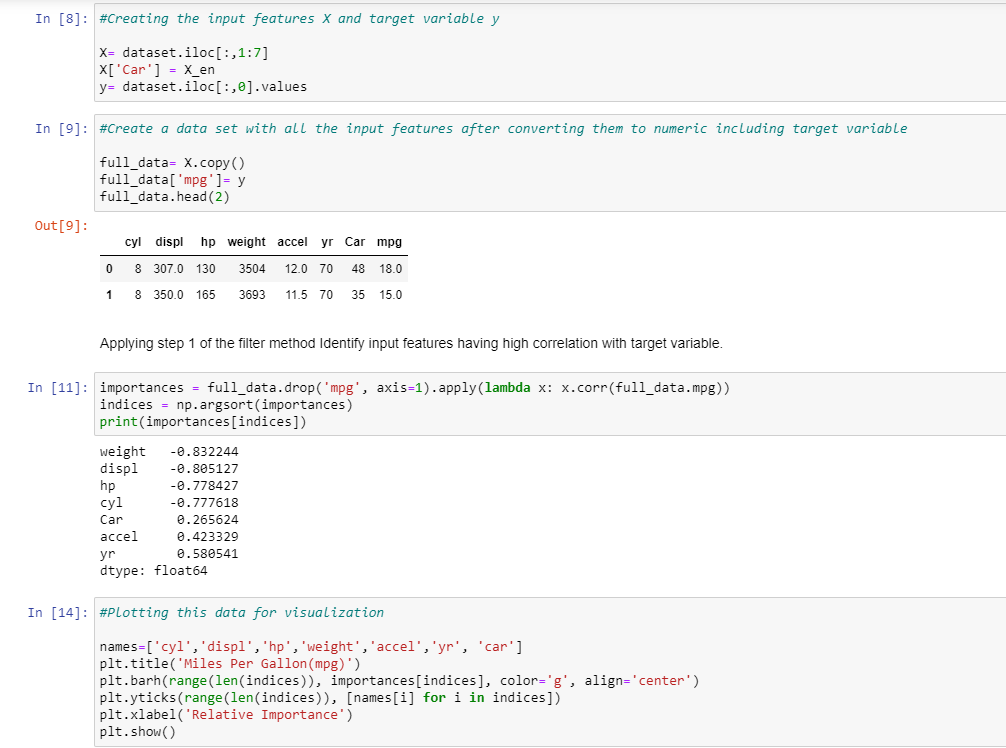
**Algorithm/Flowchart/Code/Sample Outputs**

**# Pearson’s Correlation with f\_regression function**

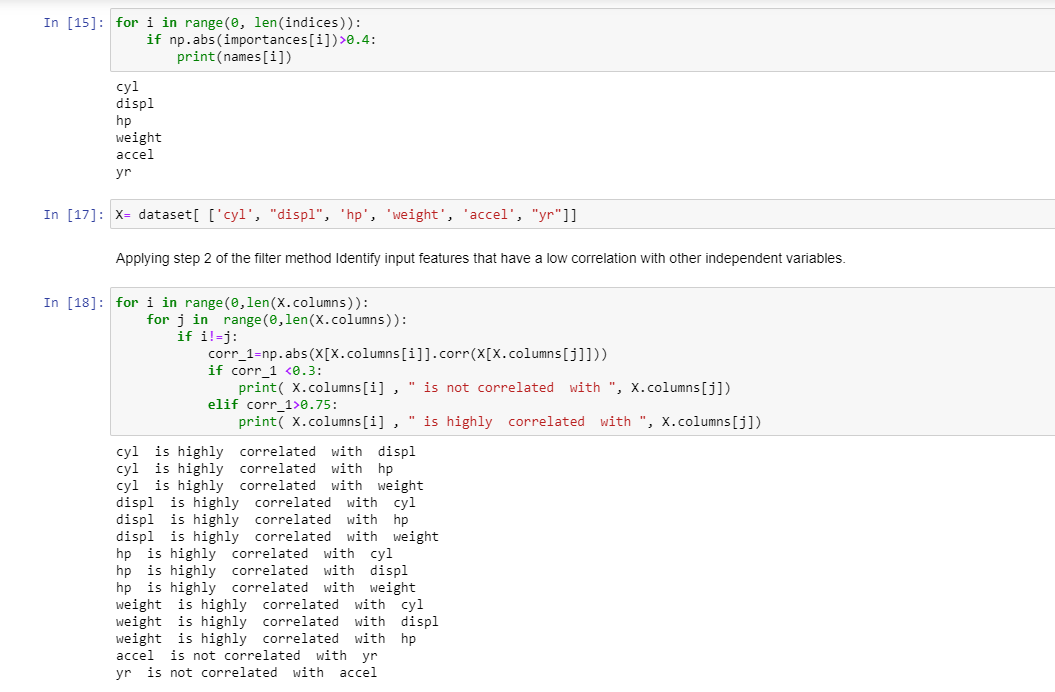
**# Creating regression dataset with make\_regression function**

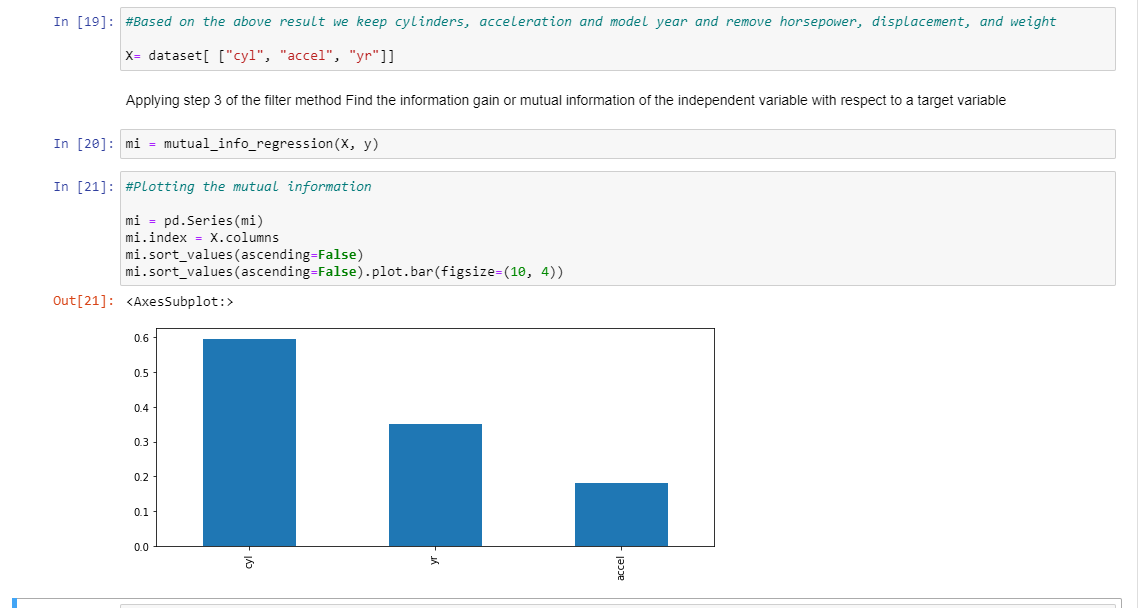












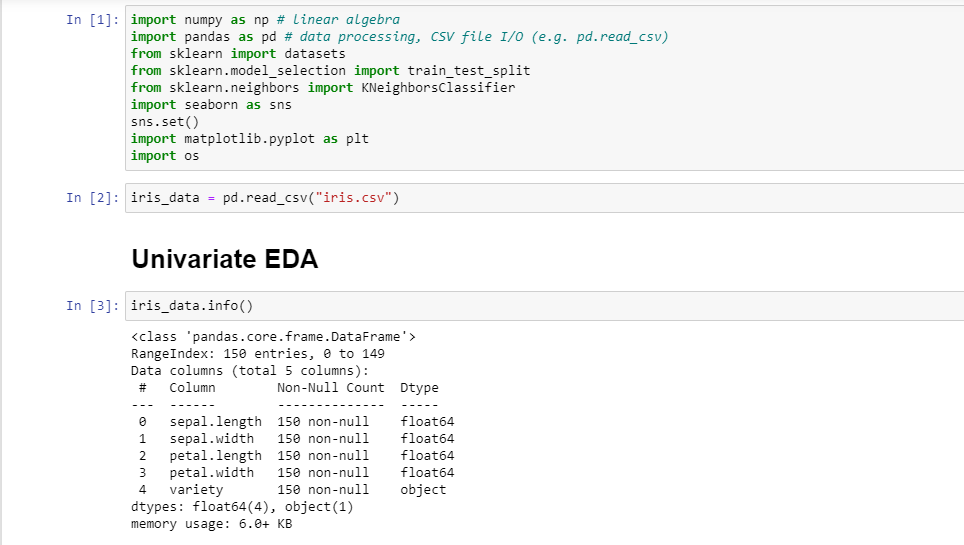
**EXPERIMENT NO. 5**

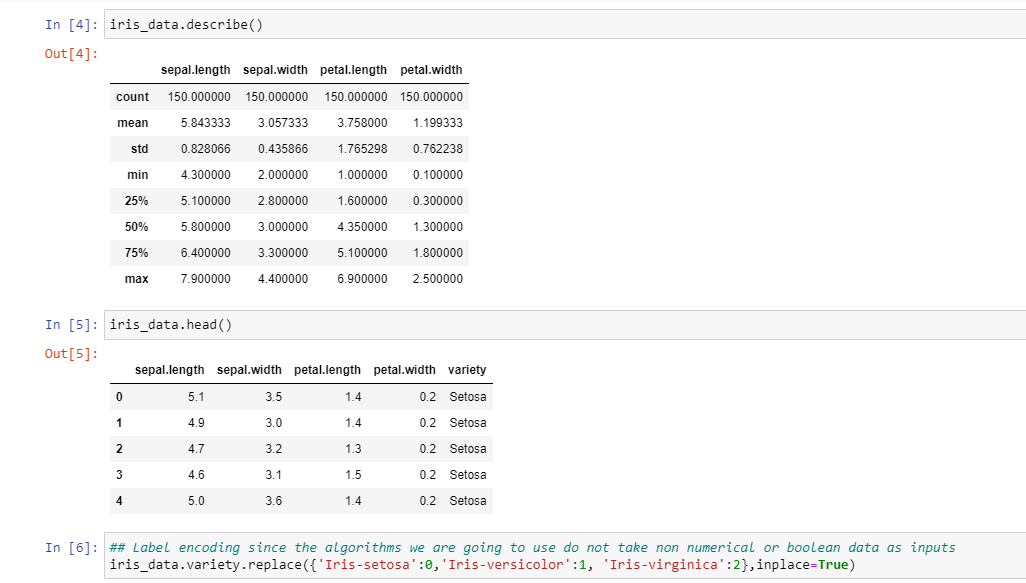
|  |
| --- |
| **Student Name and Roll Number:** Mohit Khatana 20CSU294 |
| **Semester /Section:** 5TH| AIML - B2 |
| **Link to Code:** |
| **Date: 08/09/22** |
| **Faculty Signature:** |
| **Grade:** |

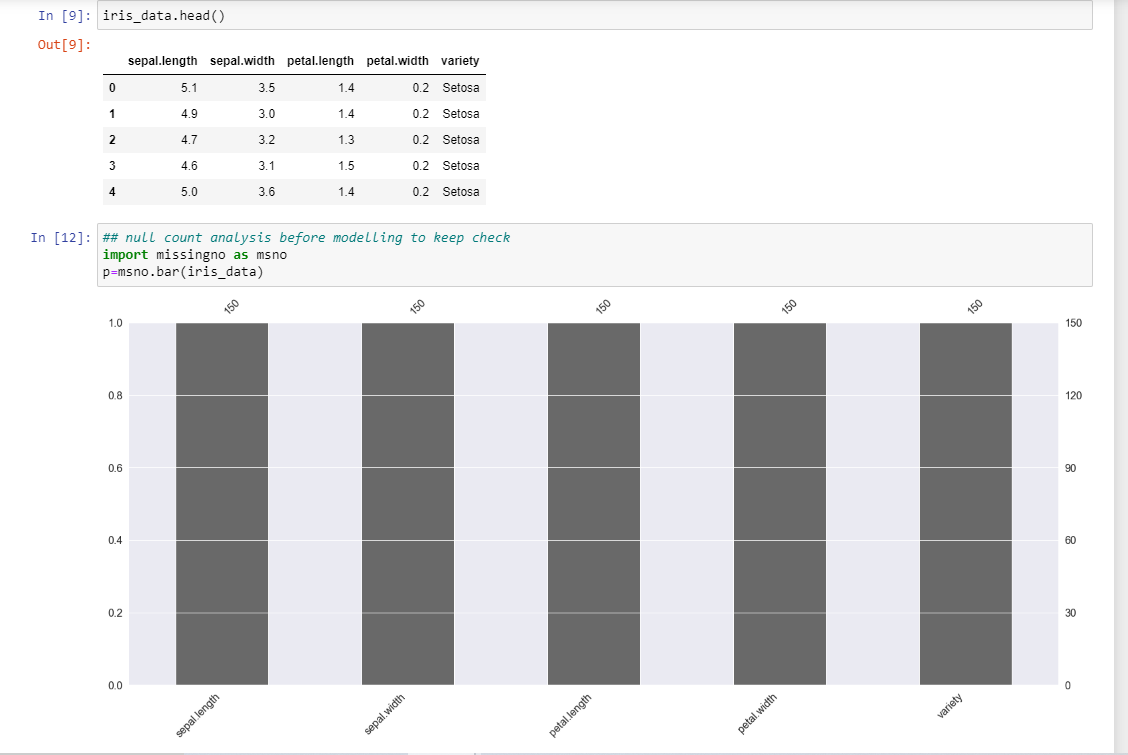
|  |
| --- |
| **Objective(s):**   * Study Dimensionality Reduction. * Understand the basic principle behind Principal Component Analysis. |
| **Outcome:**  Students will be familiarized with Dimensionality Reduction especially Principal Component  Analysis (PCA). |
| **Problem Statement:**  Reduce dimensionality of Iris dataset using Principal Component Analysis. |
| **Background Study:** Principal component analysis is a statistical technique that is used to analyze the interrelationships among a large number of variables and to explain these variables in terms of a smaller number of variables, called principal components, with a minimum loss of information. |
| **Question Bank:**   1. What is dimensionality reduction?   **Ans) Dimensionality reduction is the process of reducing the number of random variables under consideration, by obtaining a set of principal variables. It can be divided into feature selection and feature extraction.**   1. Differentiate between Feature Selection, Feature Engineering and Dimensionality Reduction.   **Ans: Feature Selection**   * **Feature selection yields a subset of features from the original set of features, which are best representatives of the data. It is an exhaustive search.** * **In text data, features might be size of characters or some global features of the text. Feature selection will keep only certain features of those.** * **Feature selection is done in the context of an optimization problem.**   **Dimension Reduction**   * **Dimensionality reduction is generic and only depends on the data and not on what you plan to do with it.** * **Assuming a classification problem you select the features that will help you classify your data better, while a dimensionality reduction algorithm is unaware of this and just projects the data into a lower dimensionality space.** * **That in turn can work quite well or not for your classification algorithm.**   **Feature Engineering**   * **Feature engineering is the process of using domain knowledge to extract new variables from raw data that make machine learning algorithms work.** * **Feature engineering is necessary to effectively use machine learning algorithms and thus build predictive models.**  1. What are principal components?   **Ans: Principal components are new variables that are constructed as linear combinations or mixtures of the initial variables.** |

**Student Work Area**

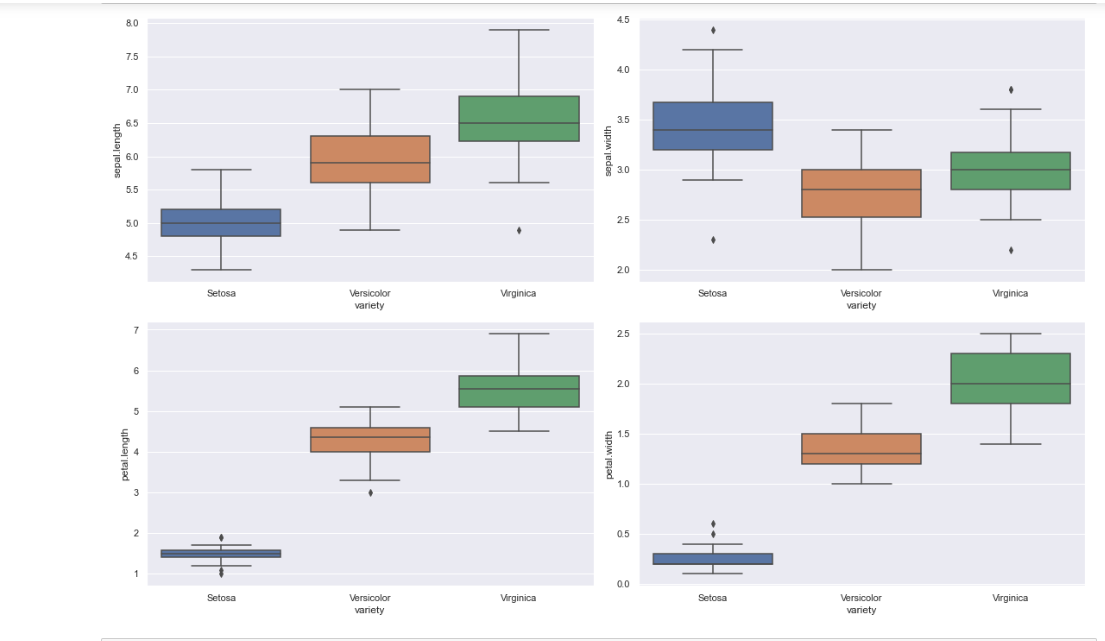
**Algorithm/Flowchart/Code/Sample Outputs**

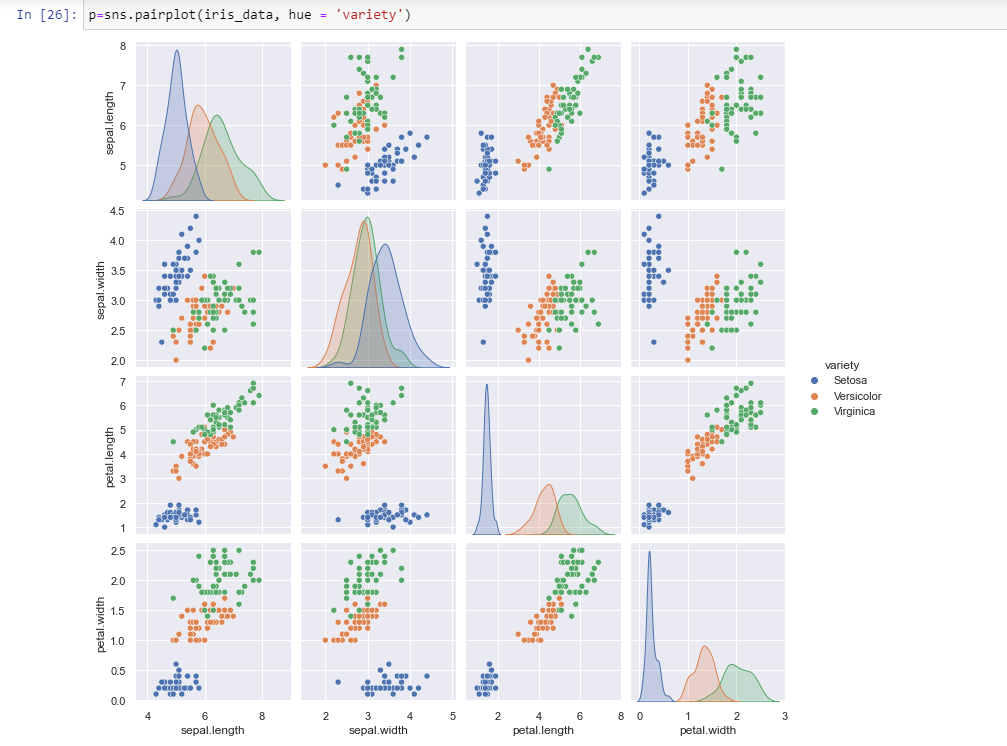






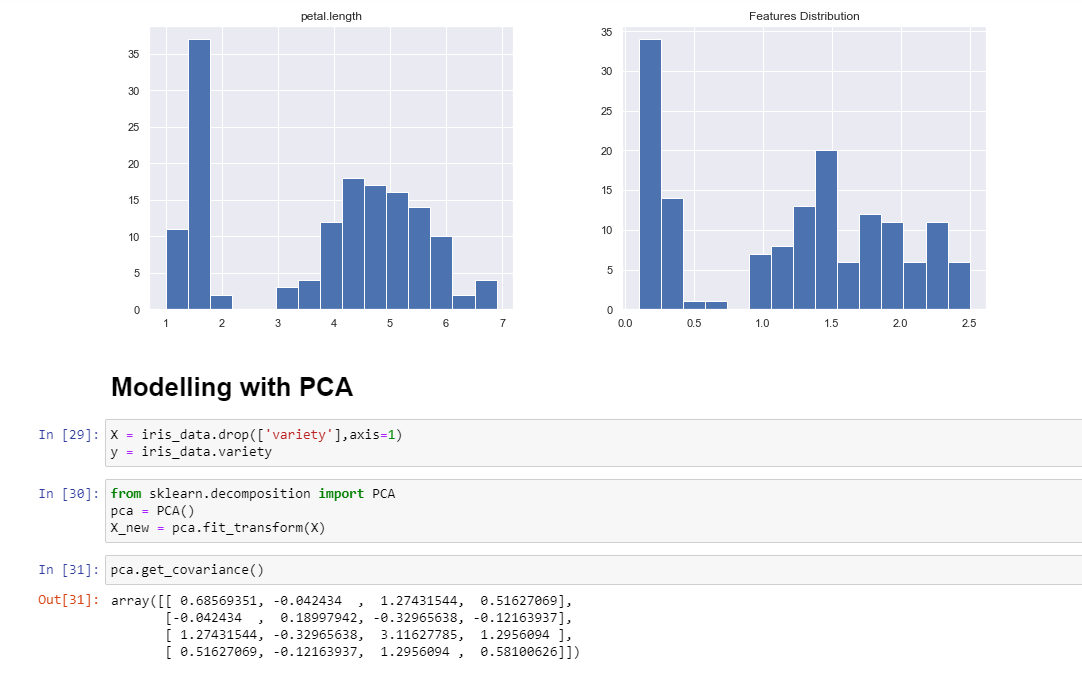


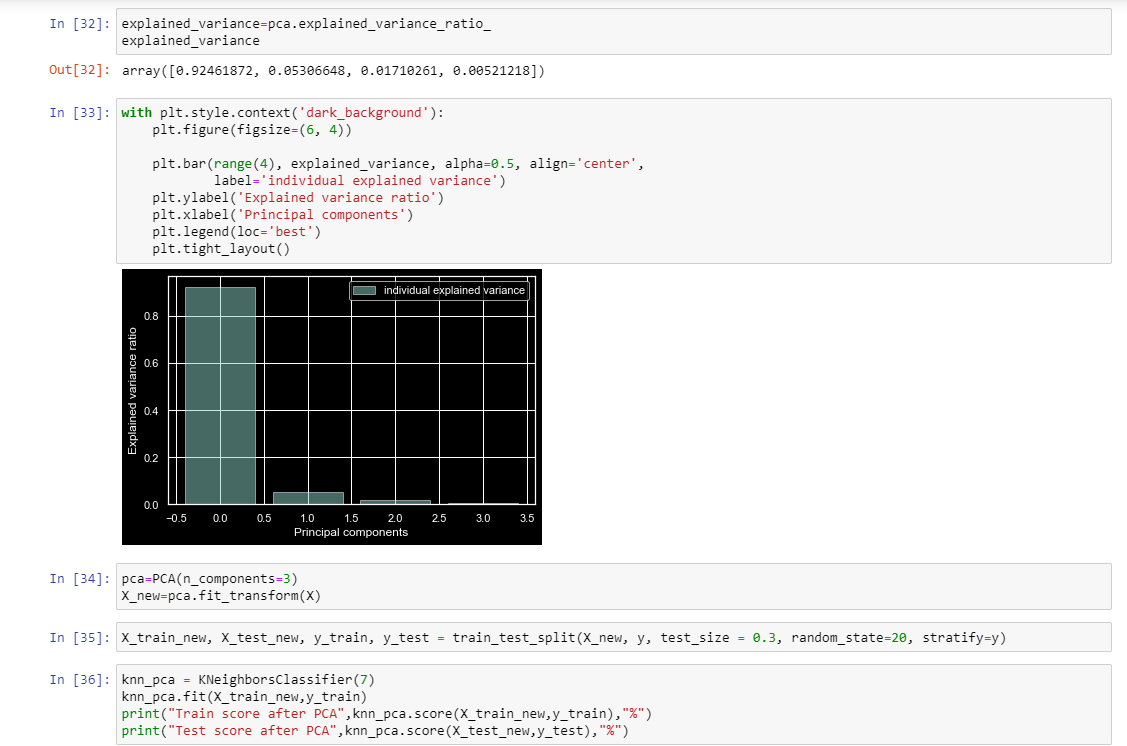


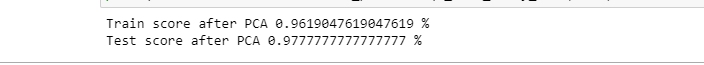












**EXPERIMENT NO. 6**

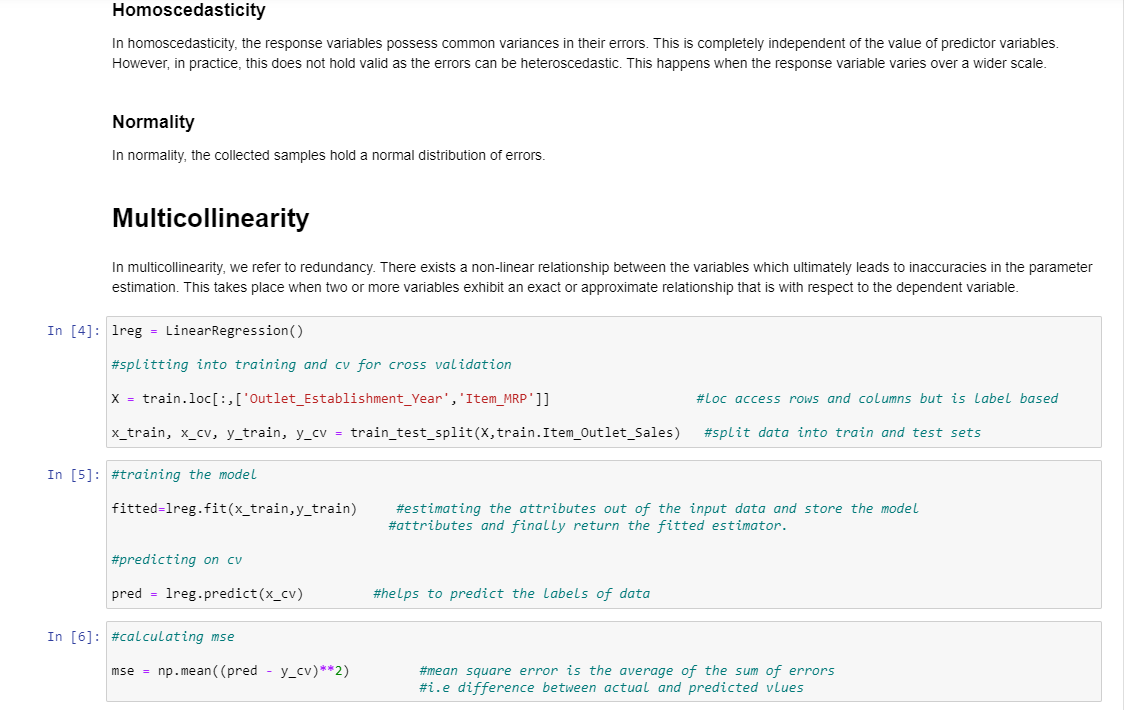
|  |
| --- |
| **Student Name and Roll Number:** Mohit Khatana 20CSU294 |
| **Semester /Section:** 5 AIML B1 |
| **Link to Code:** |
| **Date: 15/09/22** |
| **Faculty Signature:** |
| **Grade:** |

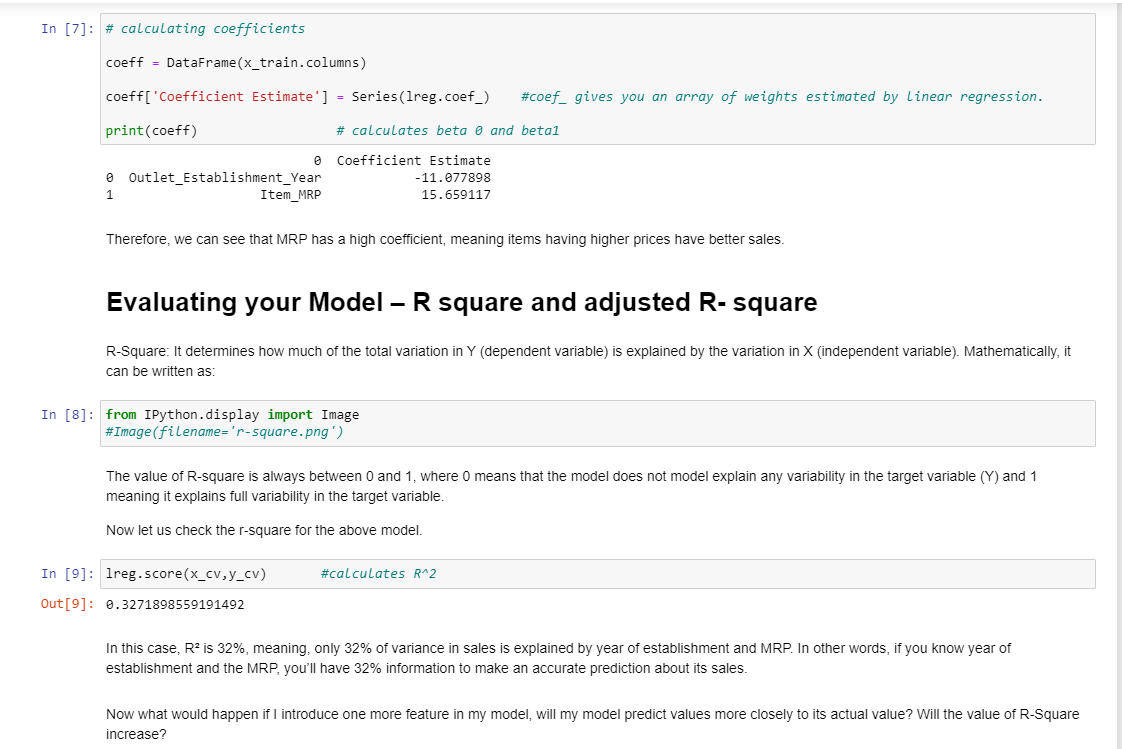
|  |
| --- |
| **Objective(s):**   * Understand Simple Linear Regression (SLR). * Study about the different performance metrics of SLR. |
| **Outcome:**  Student will be familiarized with regression problems and SLR as a solution to single feature problem. |
| **Problem Statement:**  To apply Simple Linear Regression on the given dataset. |
| **Background Study:**  Simple linear regression isan approach for predicting a response using a single feature. It is assumed that the two variables are linearly related. Hence, we try to find a linear function that predicts the response value(y) as accurately as possible as a function of the feature or independent variable(x). |
| **Question Bank:**   1. What is a regression problem?   **Ans) The regression problem is how to model one or several dependent variables/responses, Y, by means of a set of predictor variables, X.**   1. How Simple Linear Regression (SLR) helps in solving regression problems containing an input feature and an output variable?   **Ans) In Simple Linear Regression (SLR), we will have a single input variable based on which we predict the output variable.**  **Input variables can also be termed as Independent/predictor variables, and the output variable is called the dependent variable.**  **The equation for SLR is y=βo,+β1x+ϵ, where, Y is the dependent variable, X is the predictor, βo, β1 are coefficients/parameters of the model, and Epsilon(ϵ) is a random variable called Error Term.**  **OLS(Ordinary Least Squares), Gradient Descent are the two common algorithms to find the right coefficients for the minimum sum of squared errors.**   1. What are the different performance metrics that can be used for evaluating SLR?   **Ans) In order to evaluate Regression models, the metrices are:**  **Mean Absolute Error (MAE),**  **Mean Squared Error (MSE),**  **Root Mean Squared Error (RMSE),**  **R² (R-Squared).** |

**Student Work Area**

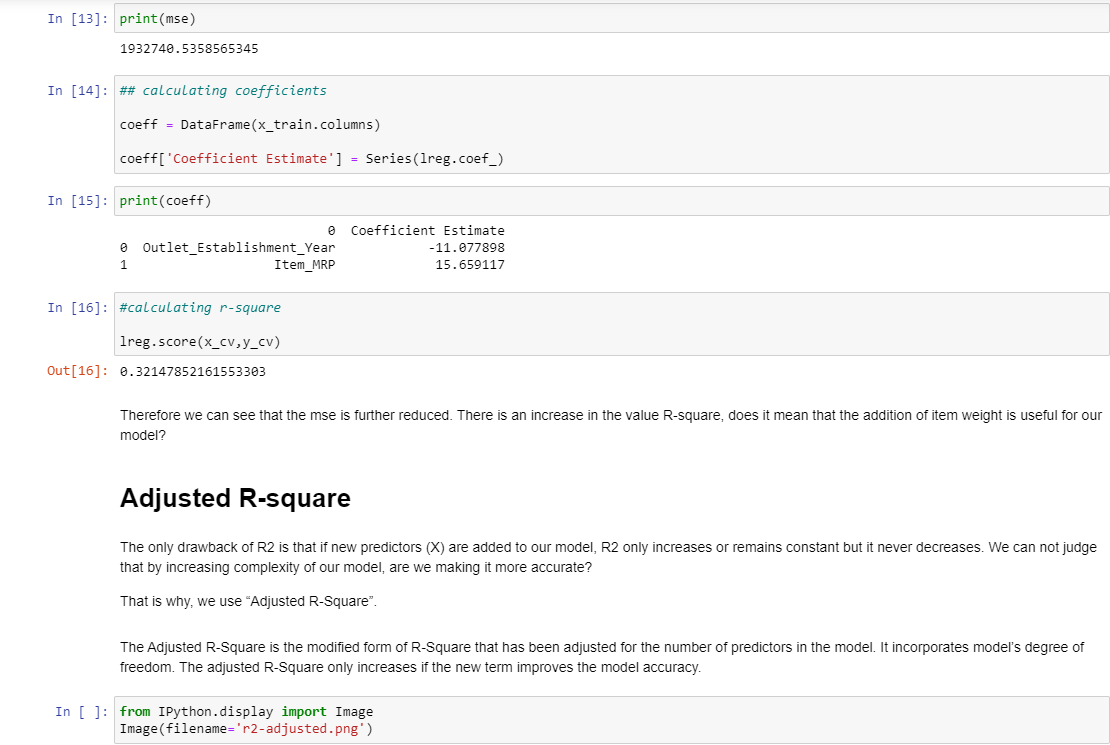
**Algorithm/Flowchart/Code/Sample Outputs**

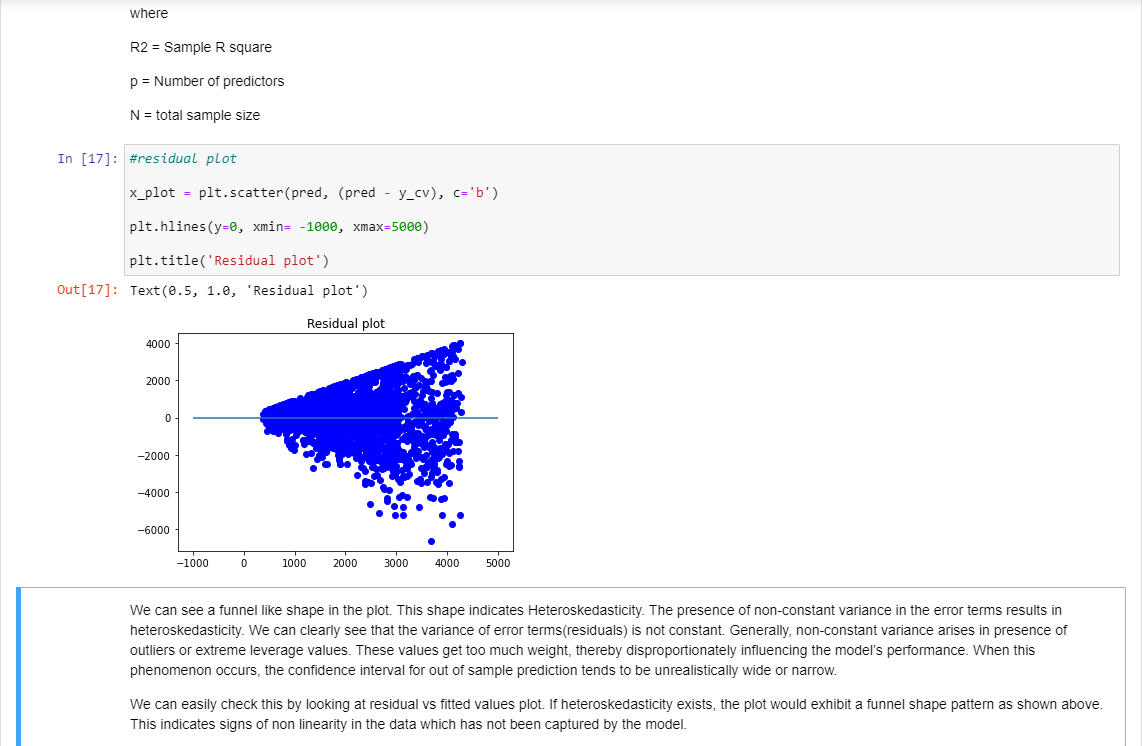












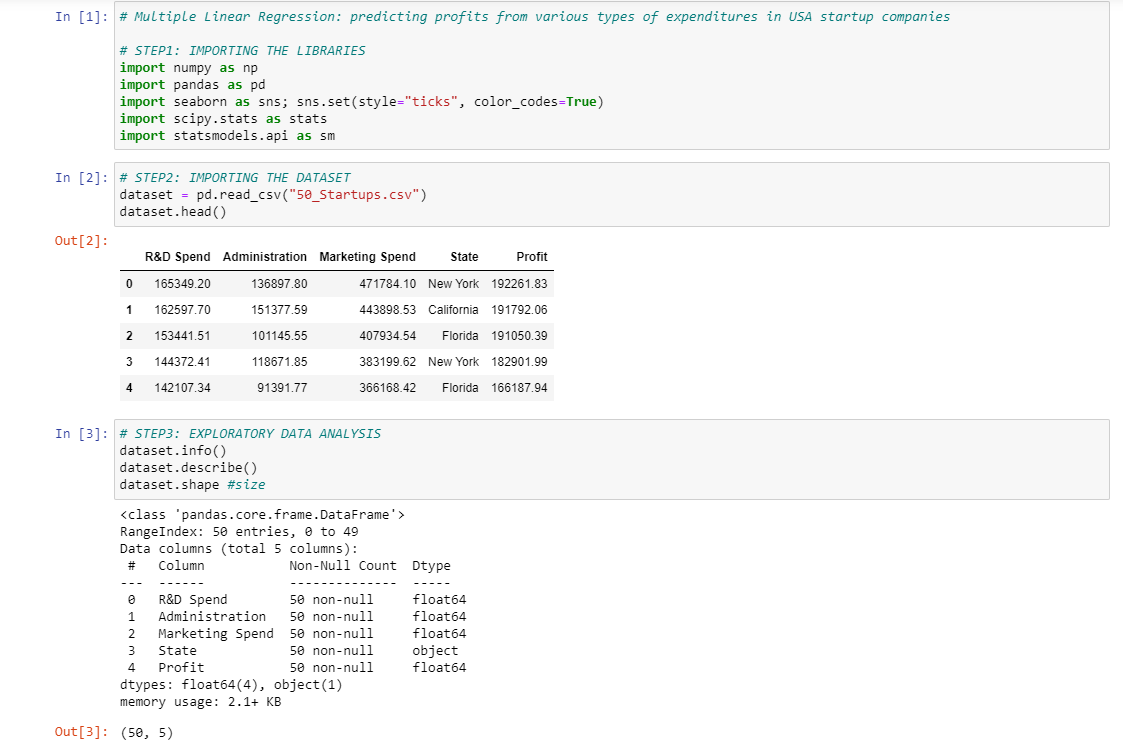
**EXPERIMENT NO. 7**

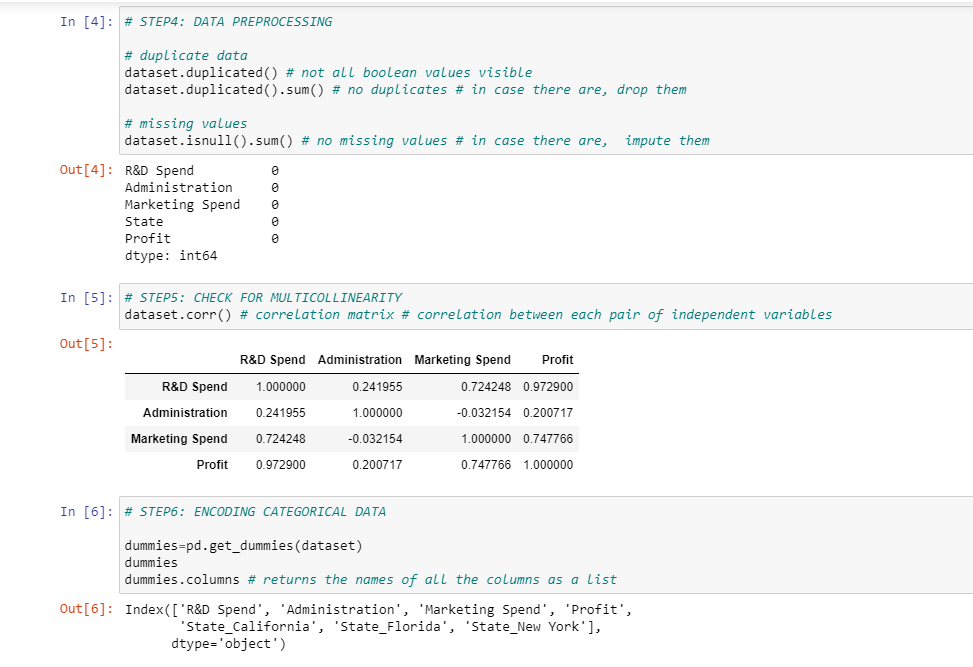
|  |
| --- |
| **Student Name and Roll Number:** Mohit Khatana 20CSU294 |
| **Semester /Section:** 5TH | AIML - B2 |
| **Link to Code:** |
| **Date: 22/09/22** |
| **Faculty Signature:** |
| **Grade:** |

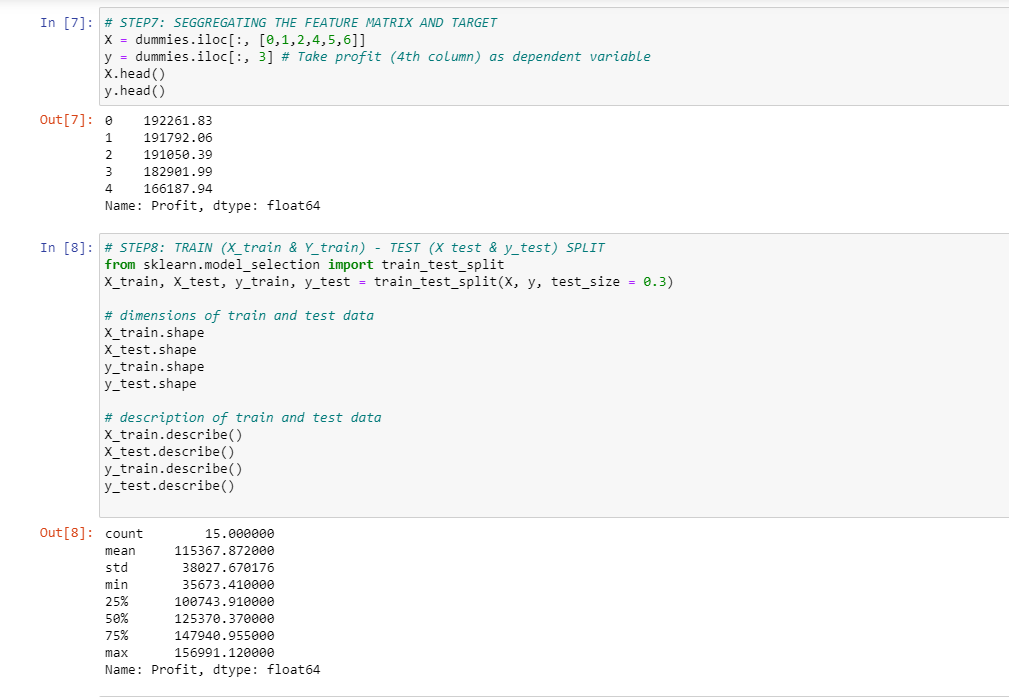
|  |
| --- |
| **Objective(s):**   * Understand mathematics behind Multiple Linear Regression (MLR). * Solving linear regression problems containing more than one independent feature using MLR. |
| **Outcome:**  Students will be familiarized with Multiple Linear Regression for solving linear regression problems. |
| **Problem Statement:**  To apply multiple linear regression on any regression dataset. |
| **Background Study:**  Multiple Linear Regressionattempts to model the relationship between two or more features and a response by fitting a linear equation to observed data. |
| **Question Bank:**   1. What is MLR?   **Ans) Multiple linear regression (MLR), also known simply as multiple regression, is a statistical technique that uses several explanatory variables to predict the outcome of a response variable.**   1. Differentiate between SLR and MLR?   **Simple linear regression: A single independent variable is used to predict the value of a dependent variable.**  **Equation: y=A+BX**  **Multiple linear regression: Two or more independent variables are used to predict the value of a dependent variable. The difference between the two is the number of independent variables.**  **Equation: y=A+BX1+CX2+DX3** |

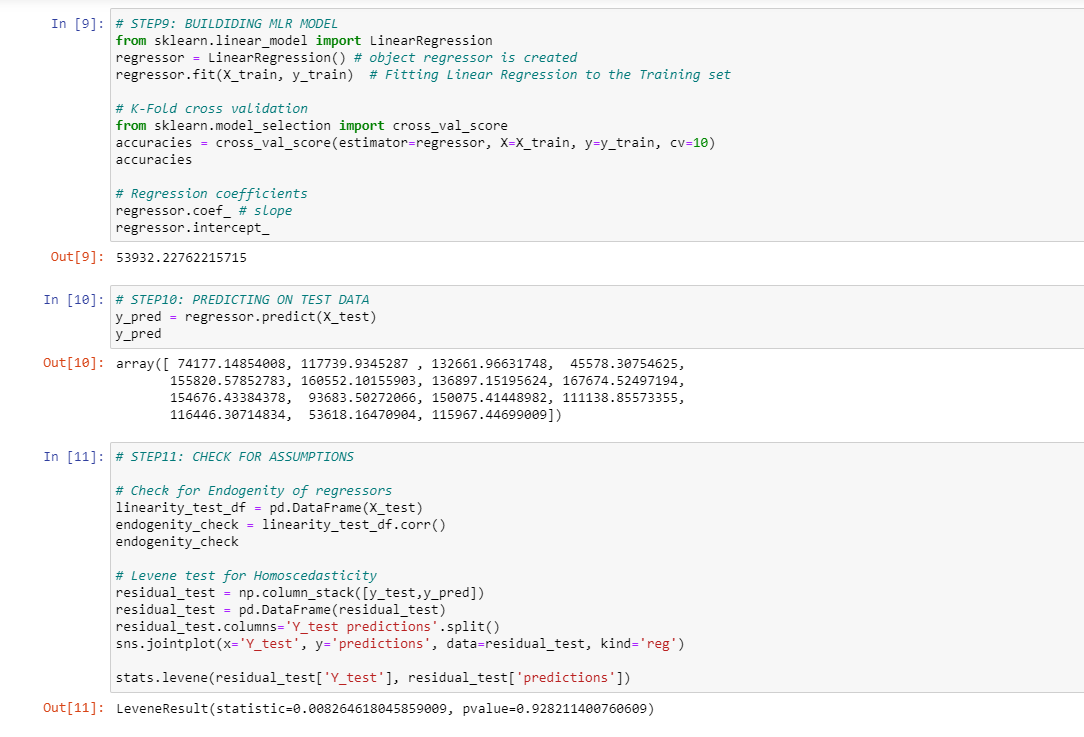
**Student Work Area**

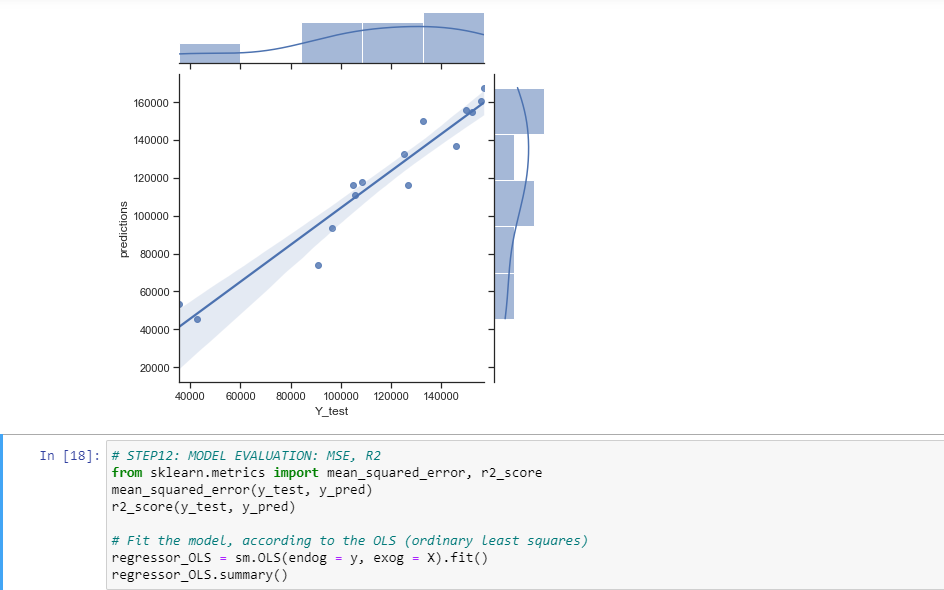
**Algorithm/Flowchart/Code/Sample Outputs**

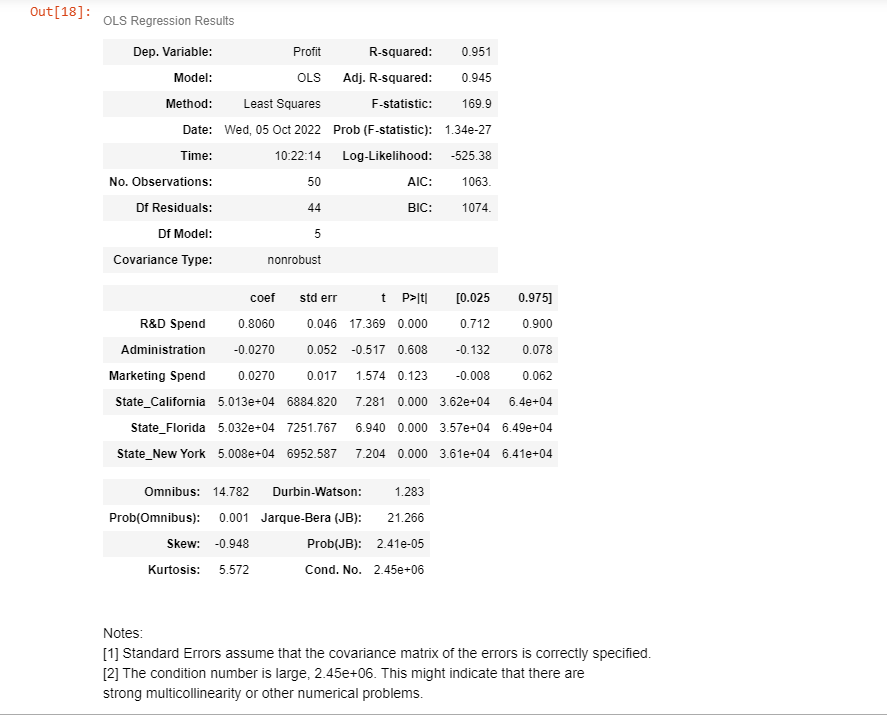












**EXPERIMENT NO. 8**

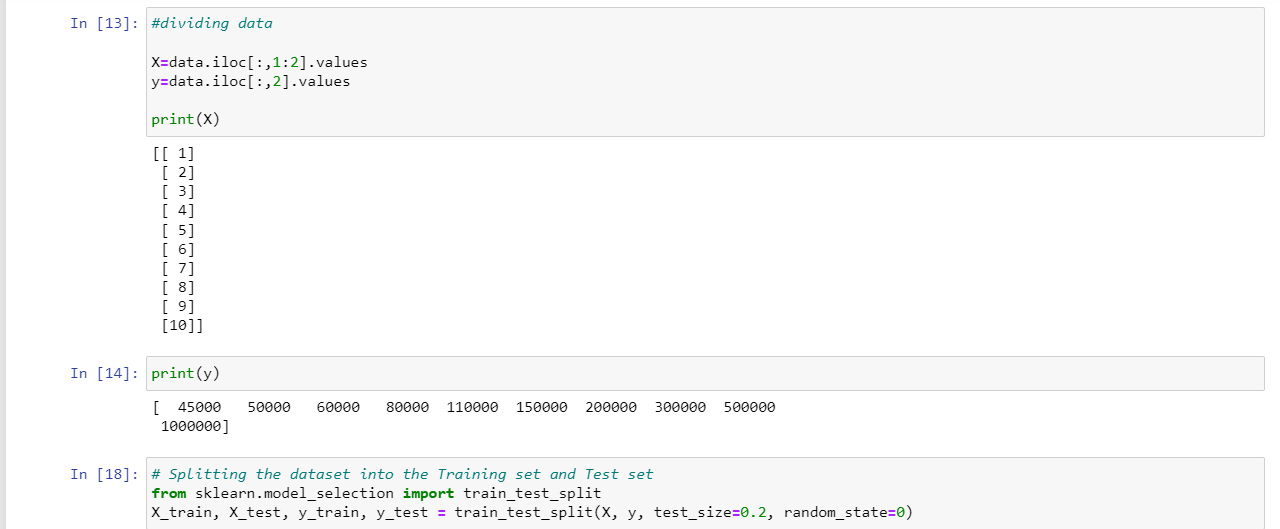
|  |
| --- |
| **Student Name and Roll Number:** Mohit Khatana 20CSU294 |
| **Semester /Section:** 5th | AIML-B |
| **Link to Code:** |
| **Date: 13/10/2022** |
| **Faculty Signature:** |
| **Grade:** |

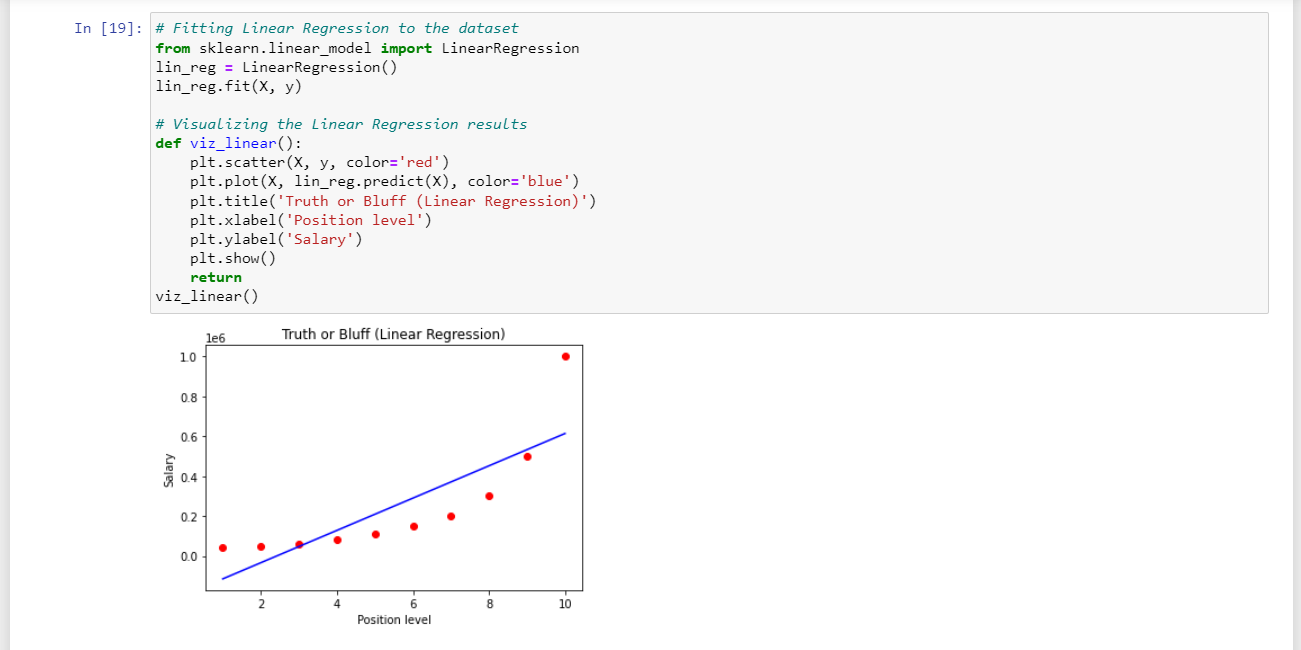
|  |
| --- |
| **Objective(s):**   * Study and understand about Polynomial Regression on non-linear regression data. * Study the mathematics behind Polynomial Regression. |
| **Outcome:**  Students will be familiarized with handling of regression data having non-linear relationship between input and output. |
| **Problem Statement:**  To apply Polynomial Linear Regression on the given dataset. |
| **Background Study:**  **Polynomial Regression**is a form of linear regression in which the relationship between the independent variable x and dependent variable y is modeled as an nth degree polynomial. Polynomial regression fits a nonlinear relationship between the value of x and the corresponding conditional mean of y, denoted E(y |x). |
| **Question Bank:**   1. What is non-linear relationship between input and output?   **Ans. In a nonlinear relationship, changes in the output do not change in direct proportion to changes in any of the inputs. While a linear relationship creates a straight line when plotted on a graph, a nonlinear relationship does not create a straight line but instead creates a curve.**   1. How Polynomial Regression is used to handle non-linear relationship?   **Ans. Polynomial regression is a form of Linear regression where only due to the Non-linear relationship between dependent and independent variables we add some polynomial terms to linear regression to convert it into Polynomial regression.** |

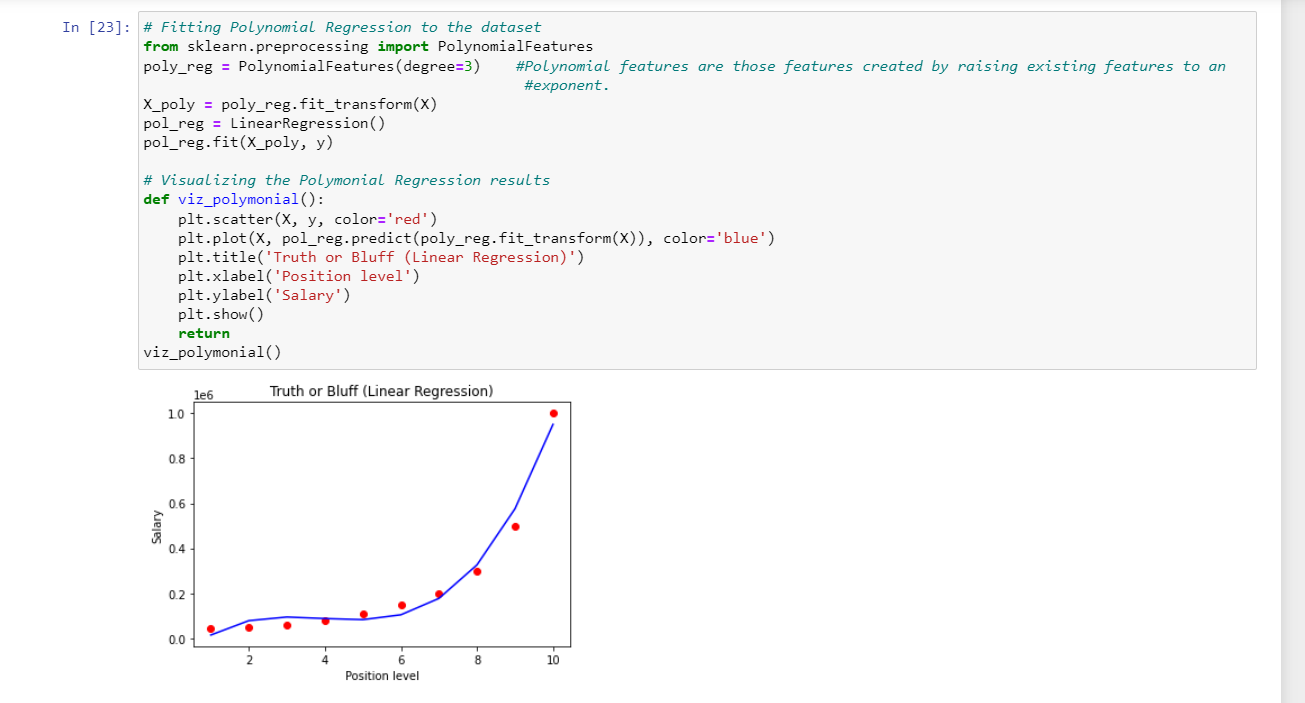
**Student Work Area**

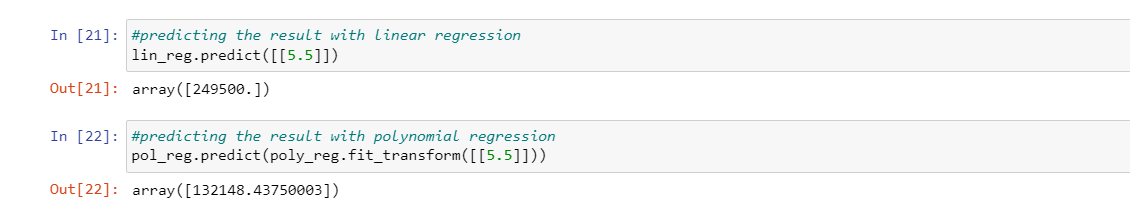
**Algorithm/Flowchart/Code/Sample Outputs**











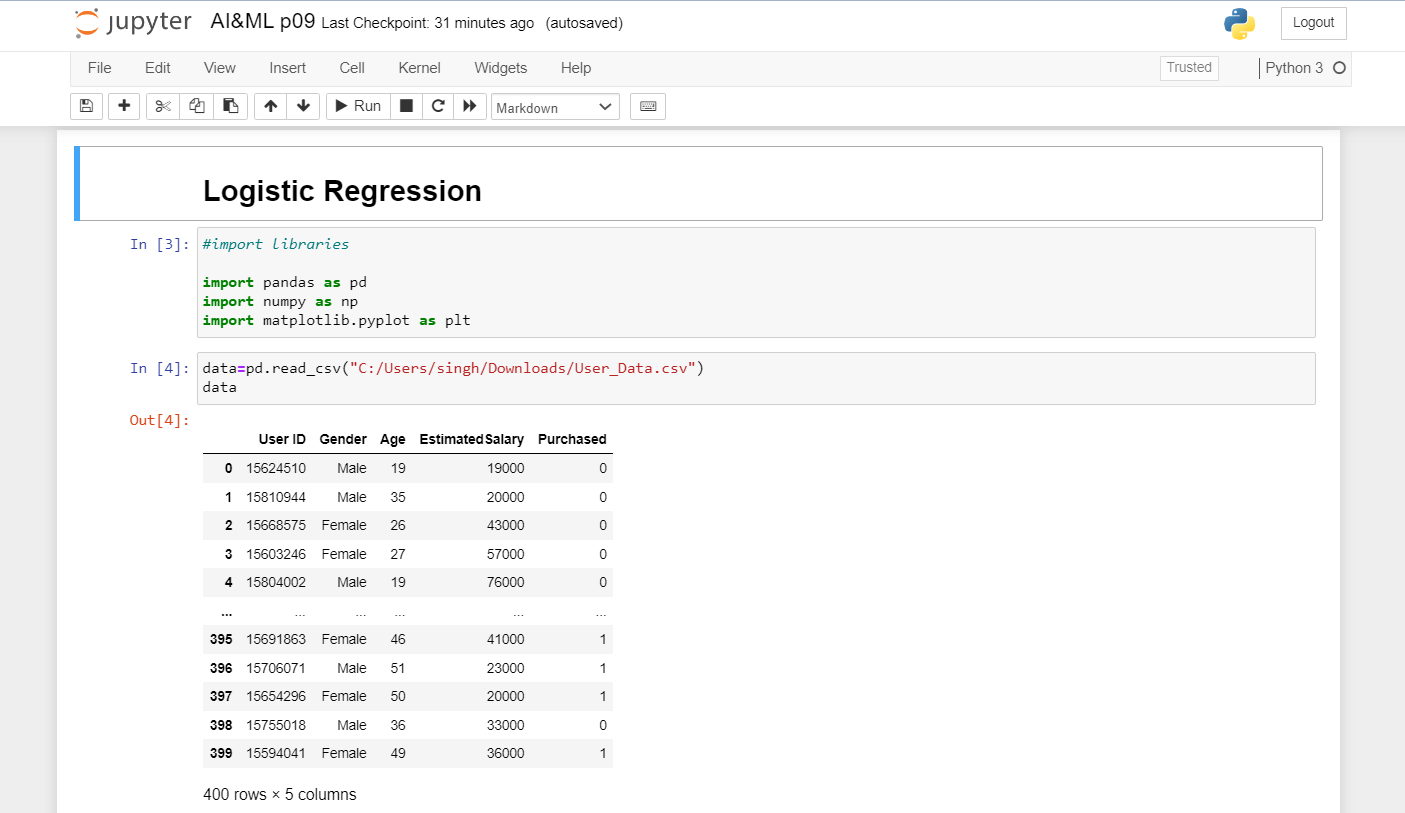
**EXPERIMENT NO. 9**

|  |
| --- |
| **Student Name and Roll Number:** Mohit Khatana 20CSU294 |
| **Semester /Section:**5 AIML B1 |
| **Link to Code:** |
| **Date: 20/10/2022** |
| **Faculty Signature:** |
| **Grade:** |

|  |
| --- |
| **Objective(s):**   * Study Logistic Regression. * How Logistic Regression is used to solve classification problems. |
| **Outcome:**  Students will be familiarized with Logistic Regression and performance metrics to calculate its performance on the given dataset. |
| **Problem Statement:**  To solve classification problems using Logistic Regression. |
| **Background Study:**  Logistic regression is a classification technique which helps to predict the probability of an outcome that can only have two values. Logistic Regression is used when the dependent variable (target) is categorical. A logistic regression produces a logistic curve, which is limited to values between 0 and 1. |
| **Question Bank:**   1. What is Logistic Regression?   Ans)Logistic regression is **an example of supervised learning.** **It is used to calculate or predict the probability of a binary (yes/no) event occurring**. An example of logistic regression could be applying machine learning to determine if a person is likely to be infected with COVID-19 or not.   1. How Logistic Regression is used for solving classification problems?   Ans)The basis of logistic regression is the logistic function, also called the **sigmoid function**, which takes in any real valued number and maps it to a value between 0 and 1.  IMG_256  Logistic regression model takes a linear equation as input and use logistic function and log odds to perform a binary classification task.   1. Why sigmoid function is used in it?   Ans)**In order to map predicted values to probabilities**, we use the Sigmoid function. The function maps any real value into another value between 0 and 1. In machine learning, we use sigmoid to map predictions to probabilities. |

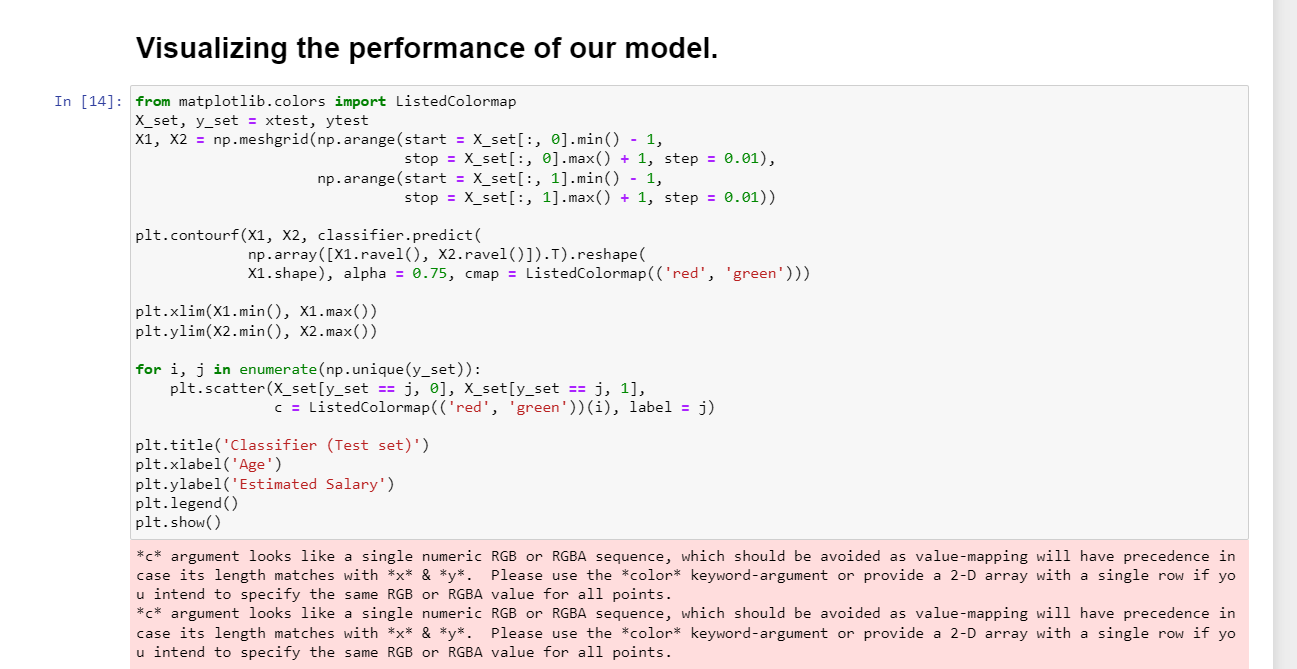
**Student Work Area**

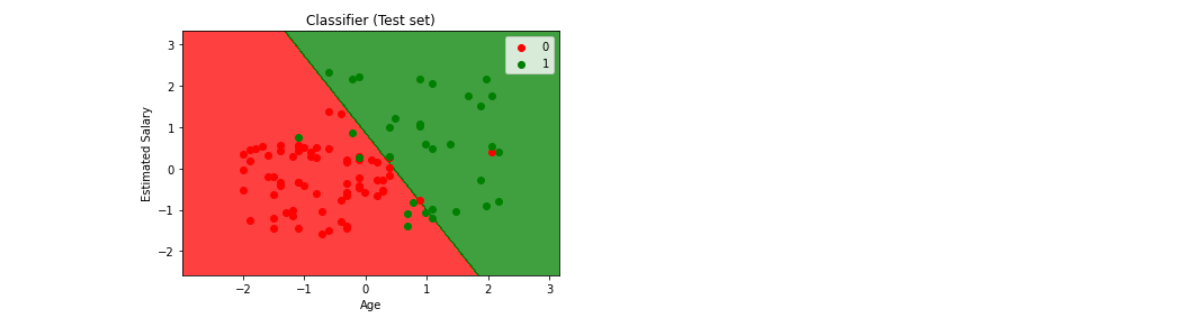
**Algorithm/Flowchart/Code/Sample Outputs**











**EXPERIMENT NO. 10**

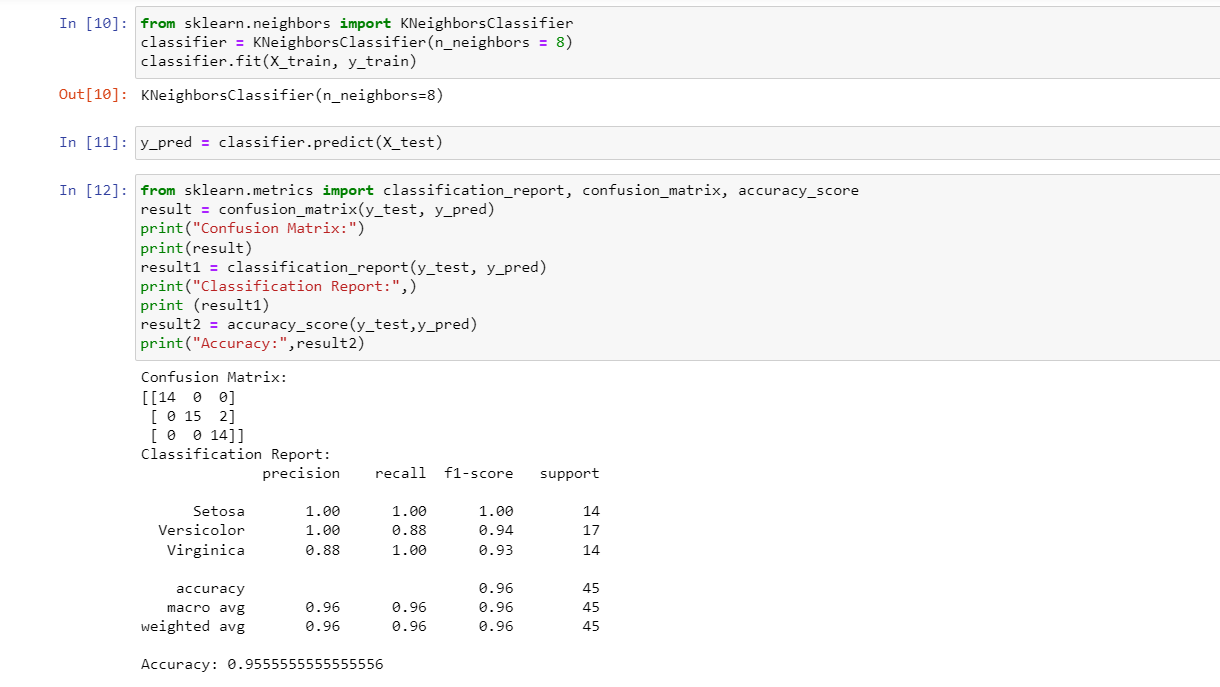
|  |
| --- |
| **Student Name and Roll Number:** Mohit Khatana 20CSU294 |
| **Semester /Section:**5 AIML B1 |
| **Link to Code:** |
| **Date:27/10/2022** |
| **Faculty Signature:** |
| **Grade:** |

|  |
| --- |
| **Objective(s):**   * Study K-Nearest Neighbour algorithm (KNN). * Understand the working principle behind KNN. |
| **Outcome:**  Students will be familiarized with classification technique using KNN. |
| **Problem Statement:**  To solve classification problems using KNN classification. |
| Background Study:  K-nearest neighbors (KNN) algorithm is a type of supervised ML algorithm which can be used for both classification as well as regression predictive problems. However, it is mainly used for classification predictive problems in industry. The following two properties would define KNN well −   * Lazy learning algorithm − KNN is a lazy learning algorithm because it does not have a specialized training phase and uses all the data for training while classification. * Non-parametric learning algorithm − KNN is also a non-parametric learning algorithm because it doesn’t assume anything about the underlying data.   K-nearest neighbors (KNN) algorithm uses ‘feature similarity’ to predict the values of new datapoints which further means that the new data point will be assigned a value based on how closely it matches the points in the training set. The most common parameter used to perform matching is Euclidean distance between the points. |
| **Question Bank:**   1. What is KNN classifier?   **Ans. the k-nearest neighbors algorithm is a non-parametric supervised learning method first developed by Evelyn Fix and Joseph Hodges in 1951, and later expanded by Thomas Cover. It is used for classification and regression. In both cases, the input consists of the k closest training examples in a data set.**   1. How KNN makes use of Euclidean distance to calculate nearest neighbor?   **Ans. To classify an unknown instance represented by some feature vectors as a point in the feature space, the k-NN classifier calculates the distances between the point and points in the training data set. Usually, the Euclidean distance is used as the distance metric.**   1. What are the other distances that can be used for nearest neighbor?   **Ans. Specifically, four different distance functions, which are Euclidean distance, cosine similarity measure, Minkowsky, correlation, and Chi square, are used in the k-NN classifier respectively.**   1. What are the various performance metrics used for classification problems?   **Ans. There are many ways for measuring classification performance. Accuracy, confusion matrix, log-loss, and AUC-ROC are some of the most popular metrics. Precision-recall is a widely used metrics for classification problems.** |

**Student Work Area**

**Algorithm/Flowchart/Code/Sample Outputs**





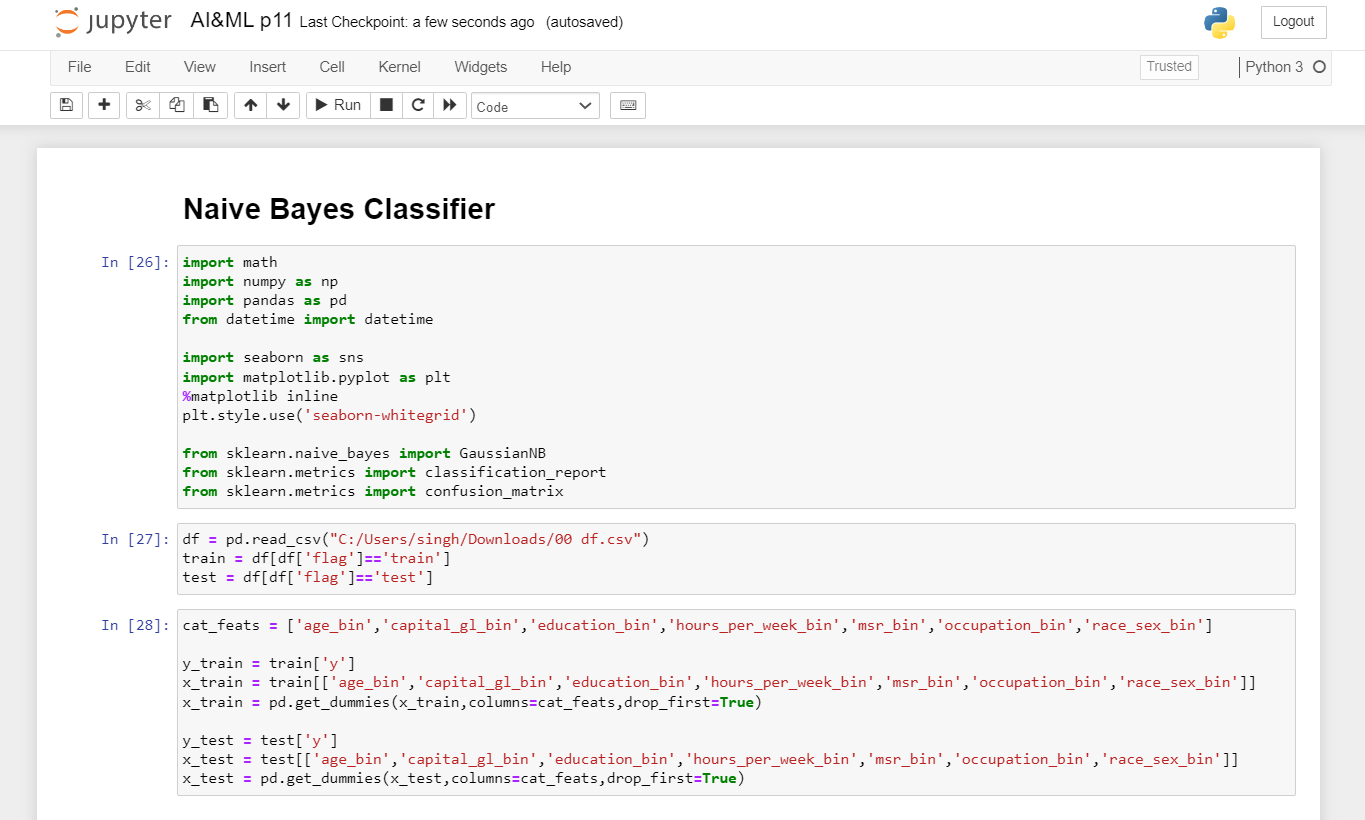
**EXPERIMENT NO. 11**

|  |
| --- |
| **Student Name and Roll Number:** Mohit Khatana 20CSU294 |
| **Semester /Section:** 5th | AIML-B2 |
| **Link to Code:** |
| **Date: 03/11/2022** |
| **Faculty Signature:** |
| **Grade:** |

|  |
| --- |
| **Objective(s):**   * Understand and study Naïve Bayes (NB) Classifier. * Understand Naïve Bayes theorem behind it. |
| **Outcome:**  Students will be familiarized with NB classification technique. |
| **Problem Statement:**  To solve classification problems using Naïve Bayes. |
| **Background Study:**  Naïve Bayes Classifier is a probabilistic classifier and is based on Bayes Theorem.  In Machine learning, a classification problem represents the selection of the Best Hypothesis given the data.  Given a new data point, we try to classify which class label this new data instance belongs to. The prior knowledge about the past data helps us in classifying the new data point. The Naïve Bayes theorem:    gives us the probability of Event A to happen given that event B has occurred. |
| **Question Bank:**   1. What is Bayes theorem?   Ans)In statistics and probability theory, the Bayes’ theorem (also known as the Bayes’ rule) is a mathematical formula used to determine the conditional probability of events. Essentially, the Bayes’ theorem describes the [probability](https://corporatefinanceinstitute.com/resources/knowledge/other/total-probability-rule/) of an event based on prior knowledge of the conditions that might be relevant to the event.  The Bayes’ theorem is expressed in the following formula:  IMG_256   1. How Naïve Bayes classifier helps for solving classification problems?   Ans) It is a classification technique based on Bayes' Theorem with an assumption of independence among predictors. In simple terms, a Naive Bayes classifier **assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature**.   1. What is the condition on features that should be fulfilled for successful application of Naïve Bayes method?   Ans)If continuous features do not have normal distribution, we should use transformation or different methods to convert it in normal distribution.  If test data set has zero frequency issue, apply smoothing techniques “Laplace Correction” to predict the class of test data set.  Remove correlated features, as the highly correlated features are voted twice in the model and it can lead to over inflating importance. |

**Student Work Area**

**Algorithm/Flowchart/Code/Sample Outputs**





**EXPERIMENT NO. 12**

|  |
| --- |
| **Student Name and Roll Number:** Mohit Khatana 20CSU294 |
| **Semester /Section: 5th / Aiml-B** |
| **Link to Code:** |
| **Date:** |
| **Faculty Signature:** |
| **Grade:** |

|  |
| --- |
| **Objective(s):**   * Understand and study Support Vector Machines (SVM). * To study how linear hyperplane is calculated to differentiate between two classes. * Basic understanding of the different variants of SVM. |
| **Outcome:**  Students will be familiarized with Support Vector Machines classifier. |
| **Problem Statement:**  To solve classification problems using SVM. |
| **Background Study:**  In machine learning, support-vector machines (SVMs, also support-vector networks) are supervised learning models with associated learning algorithms that analyze data for classification and regression analysis.  Developed at AT&T Bell Laboratories by Vladimir Vapnik with colleagues, SVMs are one of the most robust prediction methods, being based on statistical learning frameworks or VC theory proposed by Vapnik (1982, 1995) and Chervonenkis (1974).  Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier. SVM maps training examples to points in space so as tomaximise the width of the gap between the two categories. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall. |
| **Question Bank:**   1. What is SVM?   Ans. SVM or Support Vector Machine is a linear model for classification and regression problems. It can solve linear and non-linear problems and work well for many practical problems   1. What are the advantages of using SVM over other classifiers?  * SVM works relatively well when there is a clear margin of separation between classes. * SVM is more effective in high dimensional spaces and is relatively memory efficient. * SVM is effective in cases where the dimensions are greater than the number of samples.  1. What do you mean by support vectors?   Ans. Support vectors are data points that are closer to the hyperplane and influence the position and orientation of the hyperplane.  **Student Work Area**  **Algorithm/Flowchart/Code/Sample Outputs**  Graphical user interface, text, application  Description automatically generated  Graphical user interface, text, application  Description automatically generated  Text, application  Description automatically generated with medium confidence    Graphical user interface, text, application, email  Description automatically generated |

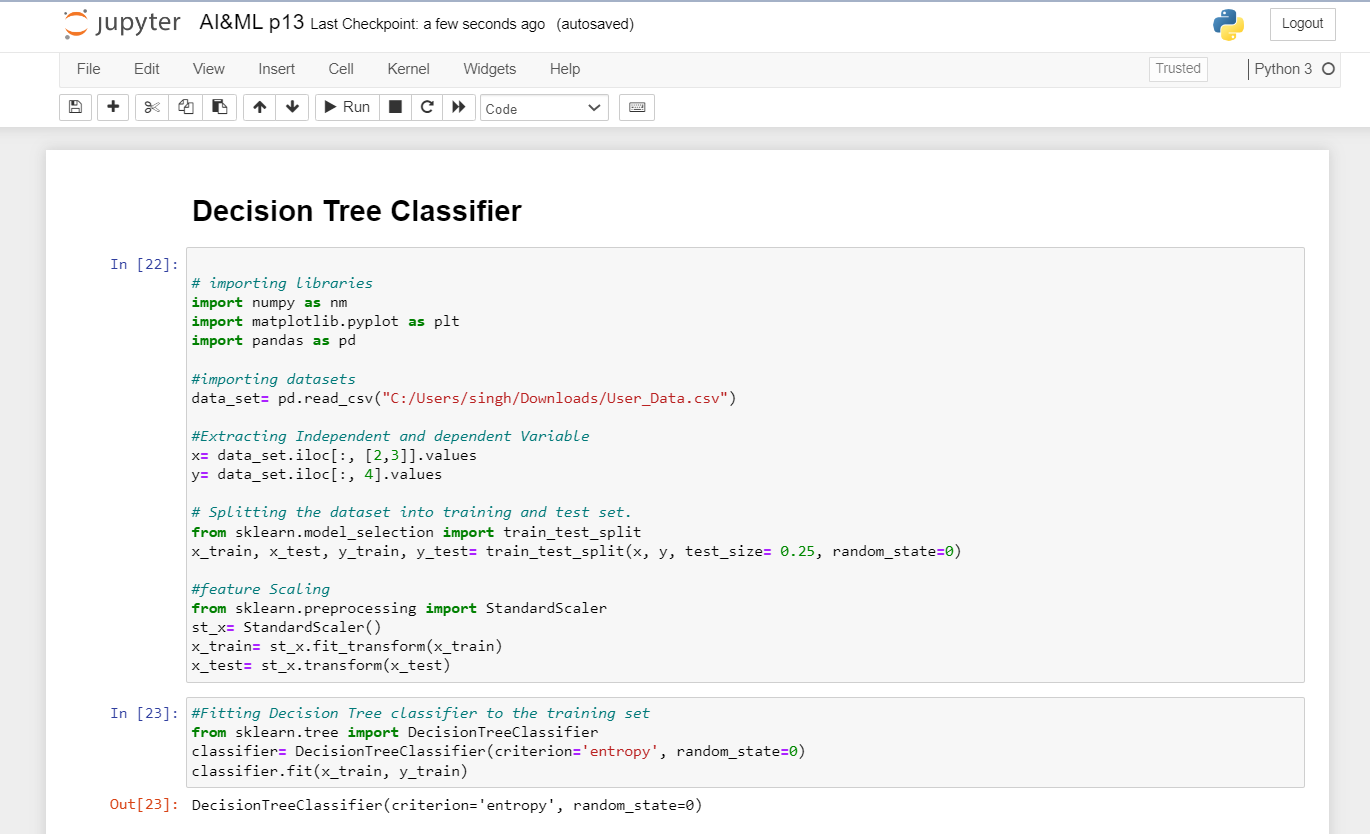
**EXPERIMENT NO. 13**

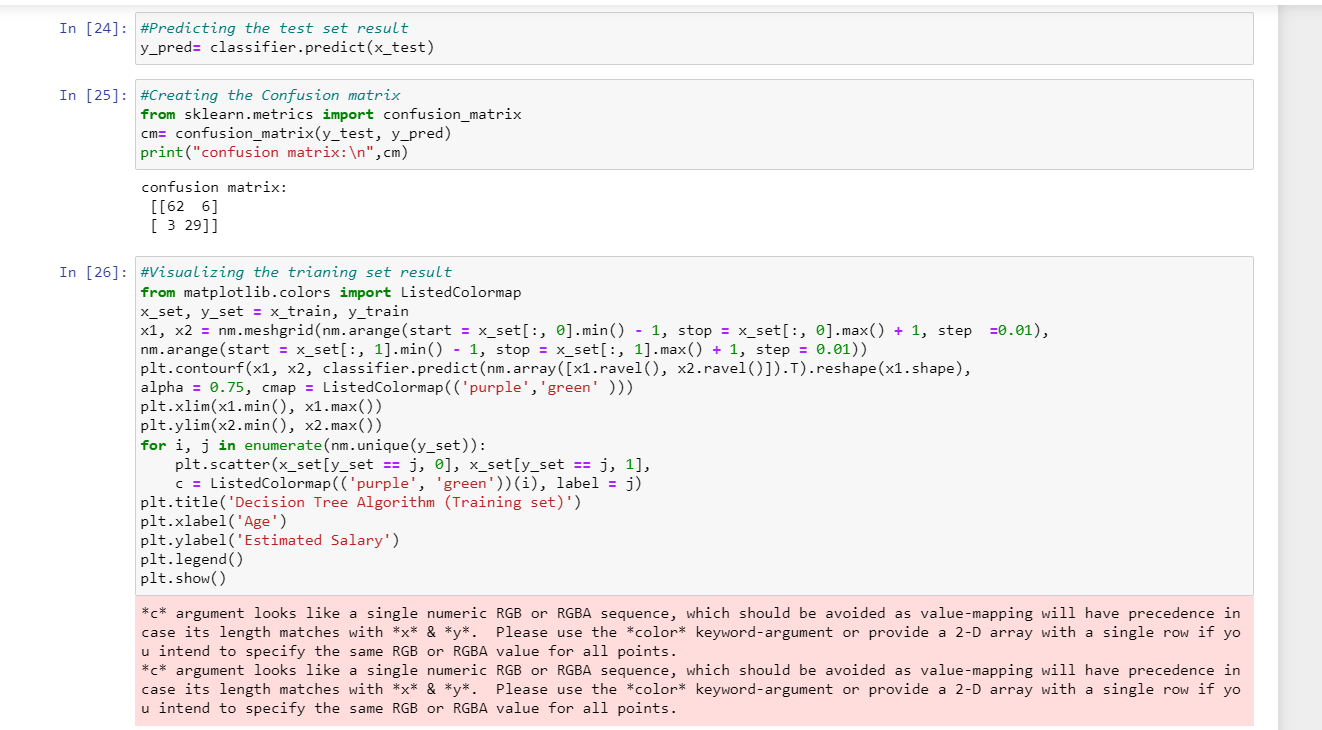
|  |
| --- |
| **Student Name and Roll Number:** Mohit Khatana 20CSU294 |
| **Semester /Section:** 5th | AIML-B2 |
| **Link to Code:** |
| **Date: 10/11/2022** |
| **Faculty Signature:** |
| **Grade:** |

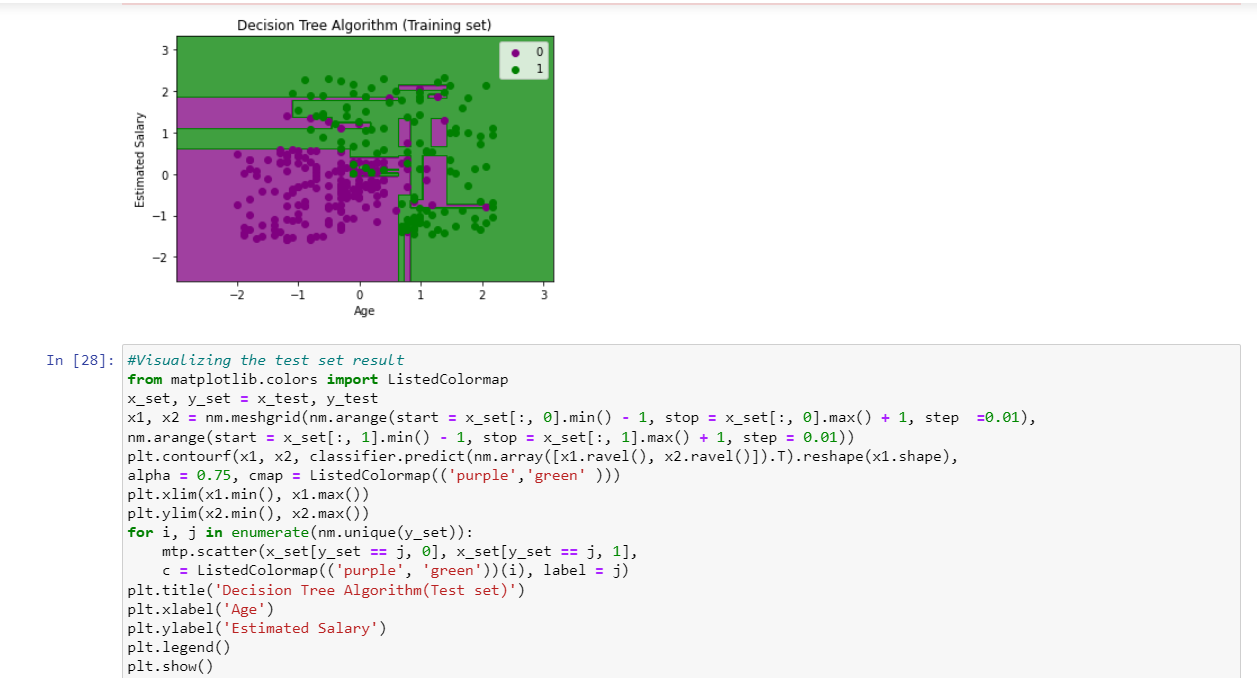
|  |
| --- |
| **Objective(s):**   * Understand and study Decision Trees for classification problems. * Study about the information gain used to create decision trees. |
| **Outcome:**  Students will be familiarized with creation of decision trees. |
| **Problem Statement:**  Apply Decision Tree classifier for solving classification problems. |
| Decision tree analysis is a predictive modelling tool that can be applied across many areas. Decision trees can be constructed by an algorithmic approach that can split the dataset in different ways based on different conditions. Decision trees are the most powerful algorithms that falls under the category of supervised algorithms.  They can be used for both classification and regression tasks. The two main entities of a tree are decision nodes, where the data is split and leaves, where we got outcome. |
| **Question Bank:**   1. What is a decision tree?   Ans)A decision tree is a non-parametric supervised learning algorithm, which is utilized for both classification and regression tasks. It has a hierarchical, tree structure, which consists of a root node, branches, internal nodes and leaf nodes.   1. How decision tree is created to solve problems?   Ans) It is a tree-structured classifier, where internal nodes represent the features of a dataset, branches represent the decision rules and each leaf node represents the outcome.  The complete process can be better understood using the below algorithm:  Step-1: Begin the tree with the root node, says S, which contains the complete dataset.  Step-2: Find the best attribute in the dataset using Attribute Selection Measure (ASM).  Step-3: Divide the S into subsets that contains possible values for the best attributes.  Step-4: Generate the decision tree node, which contains the best attribute.  Step-5: Recursively make new decision trees using the subsets of the dataset created in step -3. Continue this process until a stage is reached where you cannot further classify the nodes and called the final node as a leaf node.   1. List out the advantages and disadvantages of Decision Tree Classifiers?   Ans)IMG_256 |

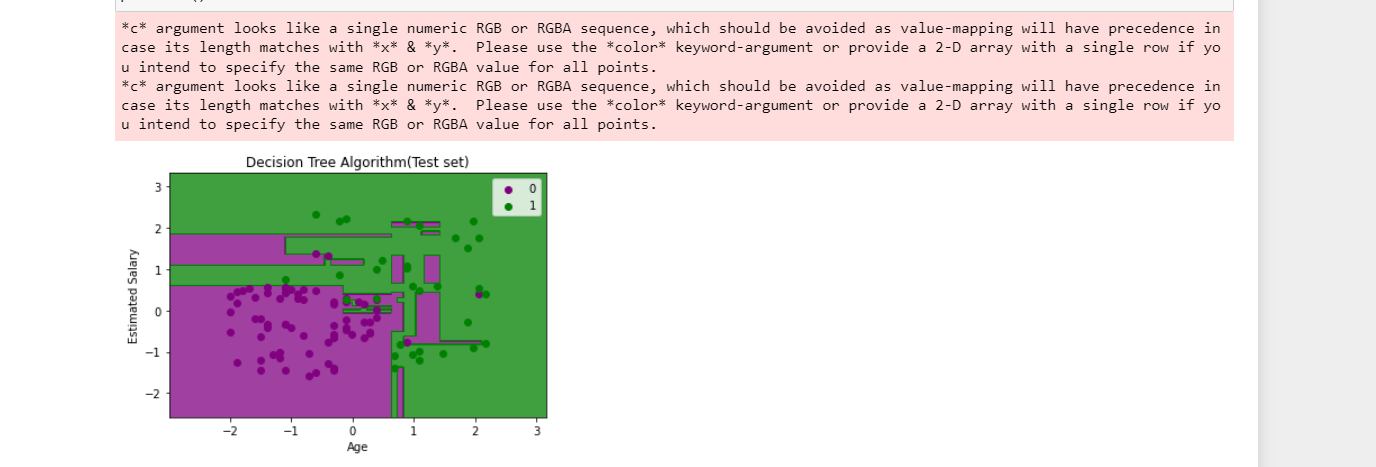
**Student Work Area**

**Algorithm/Flowchart/Code/Sample Outputs**









**EXPERIMENT NO. 14**

|  |
| --- |
| **Student Name and Roll Number:** |
| **Semester /Section:** |
| **Link to Code:** |
| **Date:** |
| **Faculty Signature:** |
| **Grade:** |

|  |
| --- |
| **Objective(s):** |
| **Outcome:** |
| **Problem Statement:** |
| **Background Study:** |
| **Question Bank:** |

**Student Work Area**

**Algorithm/Flowchart/Code/Sample Outputs**

**EXPERIMENT NO. 15**

|  |
| --- |
| **Student Name and Roll Number:** |
| **Semester /Section:** |
| **Link to Code:** |
| **Date:** |
| **Faculty Signature:** |
| **Grade:** |

|  |
| --- |
| **Objective(s):** |
| **Outcome:** |
| **Problem Statement:** |
| **Background Study:** |
| **Question Bank:** |

**Student Work Area**

**Algorithm/Flowchart/Code/Sample Outputs**