**Examination For Ai**

Archana kumari

Date-: 10-06-2024

1. What is supervised learning, and how does it differ from unsupervised learning?

Mathematical Component: Write the general form of a supervised learning model. Explain the role of the loss function L(y, y) in training the model.

Ans-:

**Supervised Learning**

* **Definition**: A type of machine learning where the model is trained on labeled data, meaning each input has a corresponding output.
* Goal: A program that performs a task as good as humans.
* TASK-well defined (the target function)
* EXPERIENCE - training data provided by a human
* PERFORMANCE-error/accuracy on the task

**Unsupervised Learning**

* **Definition**: A type of machine learning where the model is trained on data without labeled outputs.
* Goal: To find some kind of structure in the data.
* TASK-vaguely defined
* No EXPERIENCE
* NO PERFORMANCE (but, there are some evaluations metrics)

**Mathematical Component**

Model

y^​=f(X;θ)=Xθ

y^​ = predicted output.

X = input data (a matrix of features).

Θ = vector of model parameters.

choose a loss function (to fit the data)

L(y,y^​)=n1​∑i=1n​(yi​−y^​i​)2

=Value of the parameter θ

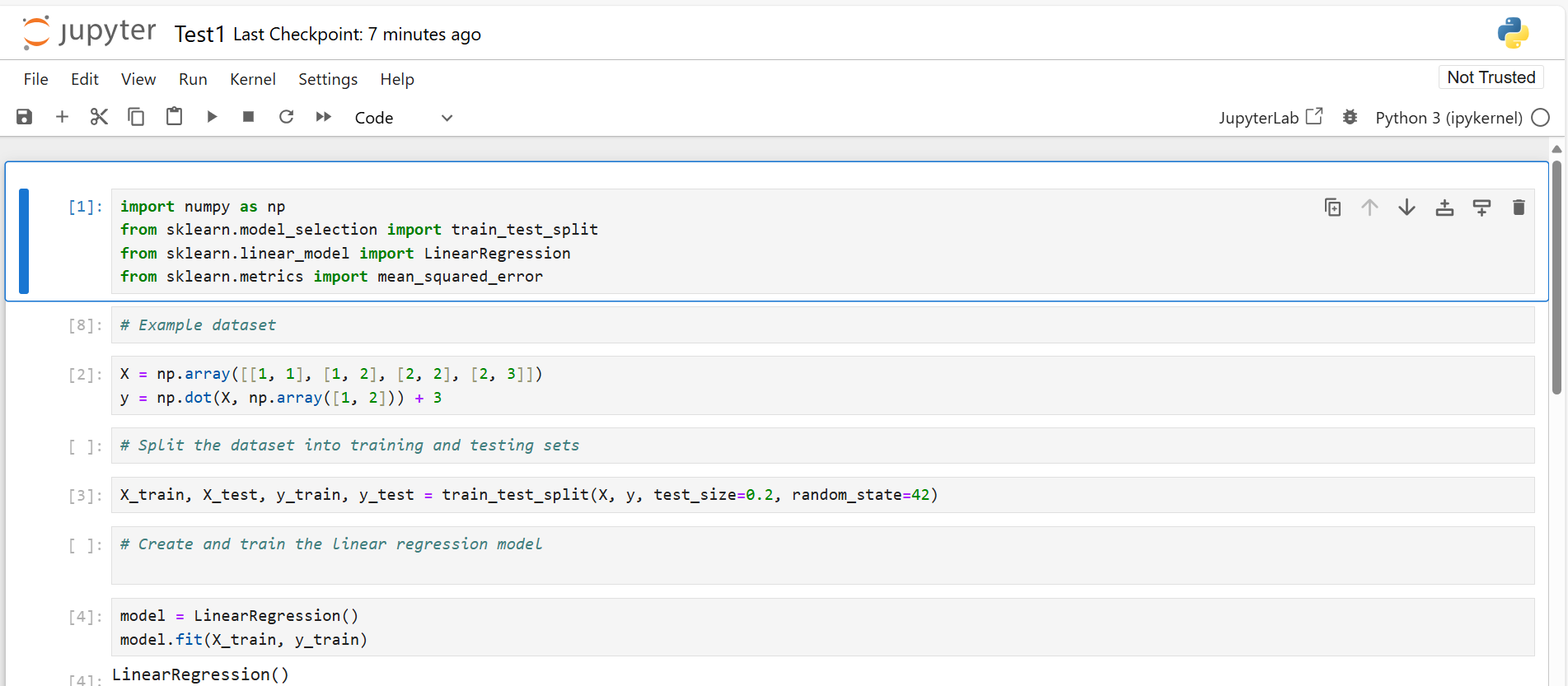
=Model to predict y^ training data X

=Compute the loss L(y,y^)

**Gradient Descent**: θ←θ−η∂L(y,y^)/∂θ

η is the learning rate.

∂L(y,y^)/∂θ​ is the gradient of the loss function with respect to the parameters.





1. Explain the assumptions underlying linear regression. How would you check if these assumptions hold for your dataset?

Mathematical Component: Given a dataset (X, y), write the equation for the linear regression model and derive the normal equations for finding the regression coefficients .

Ans-:

**Linearity**: The relationship between the input variables and the output is linear.

* **Check**: Plot observed vs. predicted values and scatter plots of each predictor vs. the response.

**Independence**: Observations are independent of each other.

* **Check**: Durbin-Watson statistic, residuals plot over time.

**Homoscedasticity**: Constant variance of residuals across all levels of the independent variables.

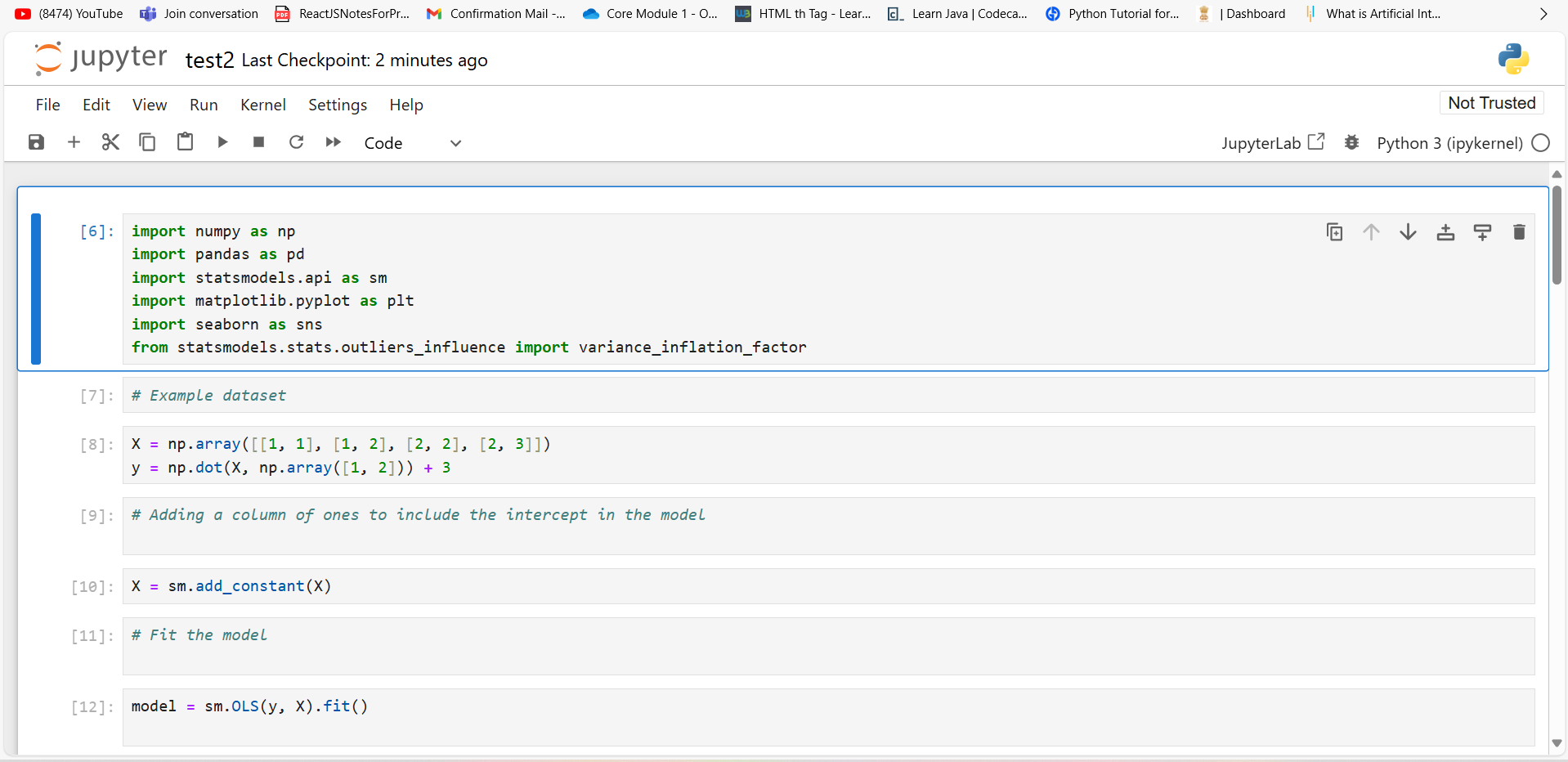
* **Check**: Plot residuals vs. predicted values, Breusch-Pagan test.

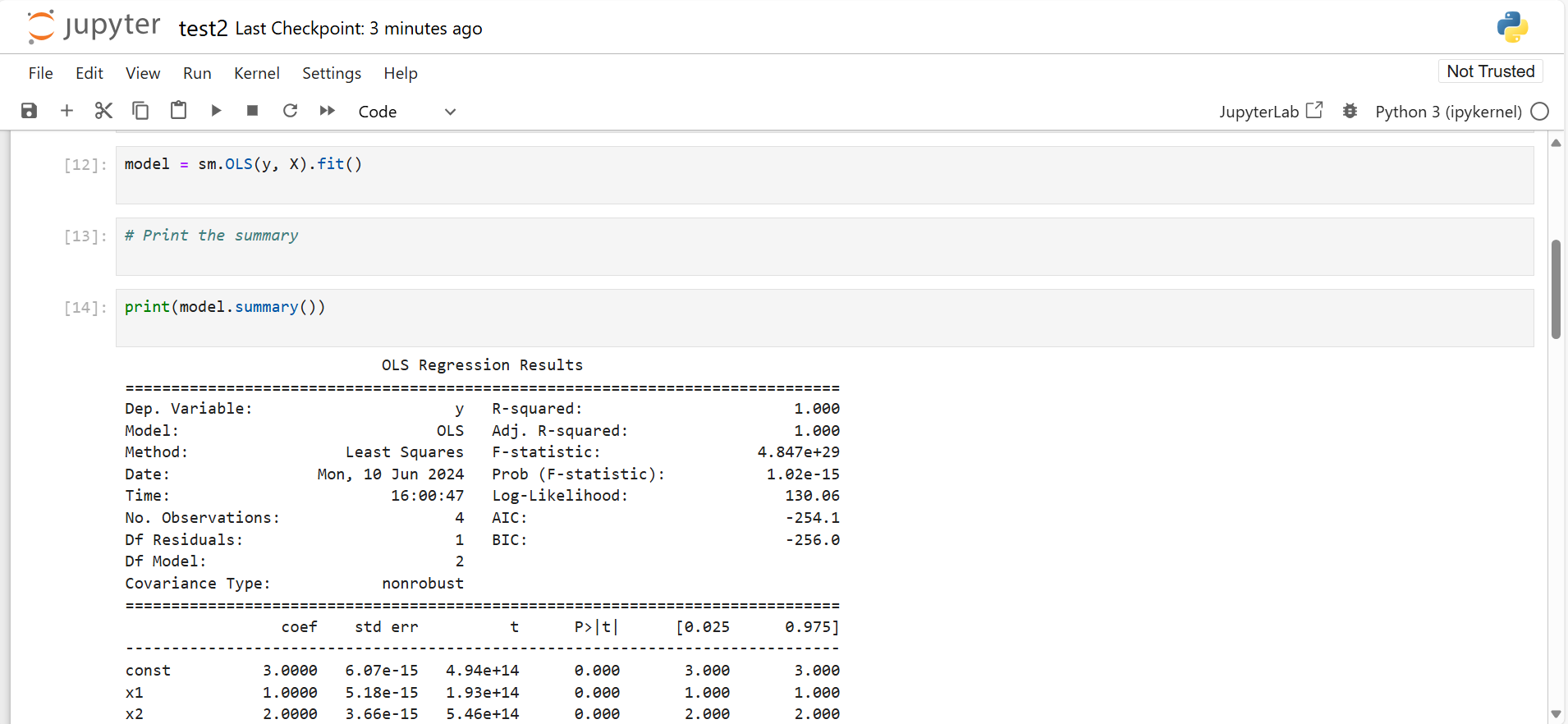
**Normality of Residuals**: Residuals are normally distributed.

* **Check**: Q-Q plot, Shapiro-Wilk test, Kolmogorov-Smirnov test.

**No Multicollinearity**: Independent variables are not highly correlated.

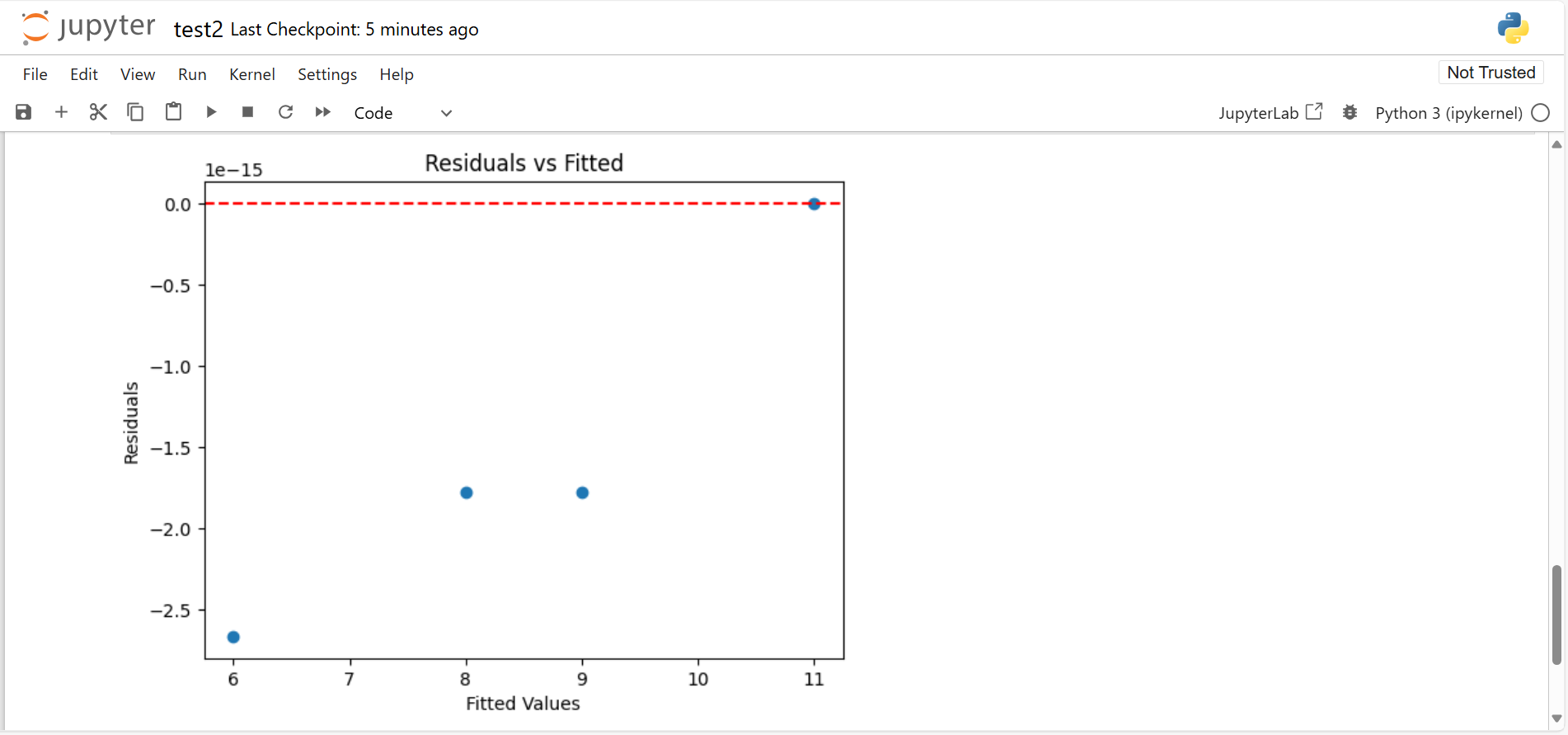
* **Check**: Variance Inflation Factor (VIF), correlation matrix.











3. Describe the difference between logistic regression and decision tree classifiers. When would you choose one over the other?

Mathematical Component: For logistic regression, write the logistic function ()= and derive the gradient of the log-likelihood function with respect to the weights.

1. **Logistic Regression**:
   * **Type**: Linear model.
   * **Purpose**: Estimates the probability of a binary outcome.
   * **Function**: Uses a logistic function to model the probability of the default class.
   * **Advantages**: Simple, interpretable, works well for linearly separable data.
   * **Use When**: Data is linearly separable, you need a probabilistic interpretation, and interpretability is important.
2. **Decision Tree**:
   * **Type**: Non-linear model.
   * **Purpose**: Classifies data by splitting it into subsets based on feature values.
   * **Function**: Creates a tree-like model of decisions and their possible consequences.
   * **Advantages**: Handles non-linear relationships, easy to visualize, captures interactions between features.
   * **Use When**: Data is not linearly separable, you need to capture complex patterns, or when interpretability and visualization are important.

**Choosing Between Them**

* **Choose Logistic Regression**:
  + When the relationship between input features and output is linear.
  + When you need a straightforward and interpretable model.
  + When you require probability estimates for the outcome.
* **Choose Decision Tree**:
  + When the data has complex, non-linear relationships.
  + When you want a model that can easily capture interactions between features.
  + When interpretability in the form of a tree structure is beneficial.

4 .

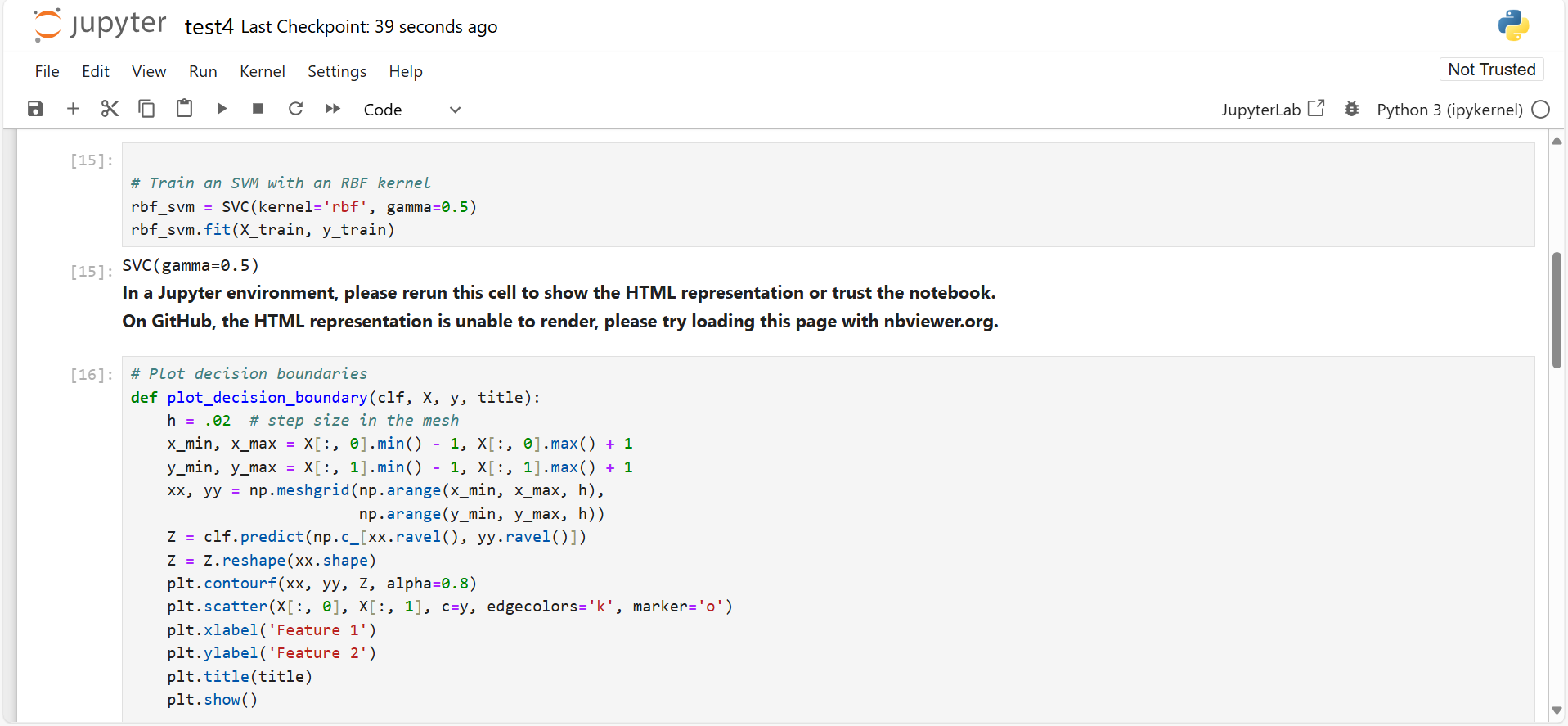
How does a Support Vector Machine (SVM) classify data? Explain the concept of the kernel trick.

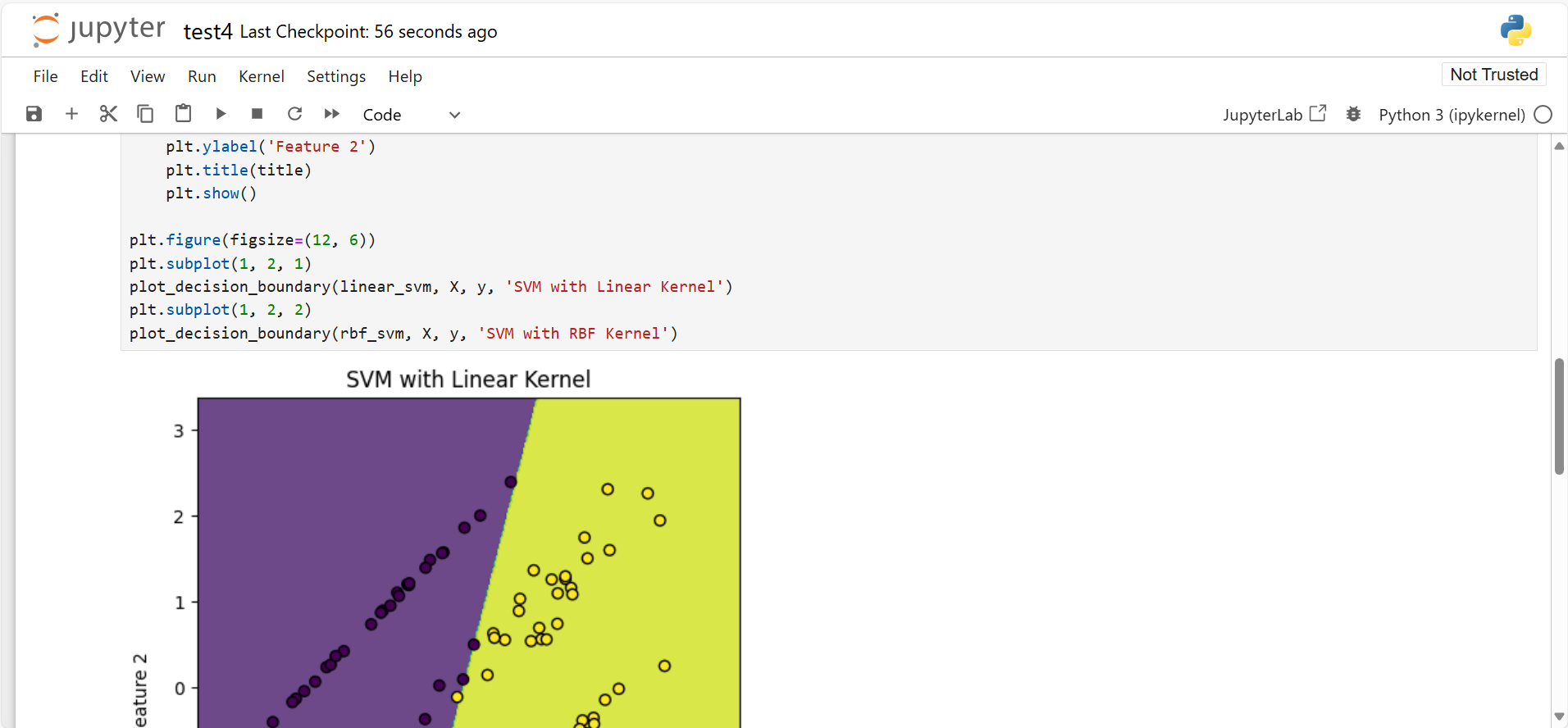
Mathematical Component: Write the optimization problem for the linear SVM and describe the constraints.

Ans-:

A Support Vector Machine (SVM) is a supervised learning algorithm that can be used for both classification and regression tasks. It works by finding the optimal hyperplane that best separates data points of different classes in a feature space.









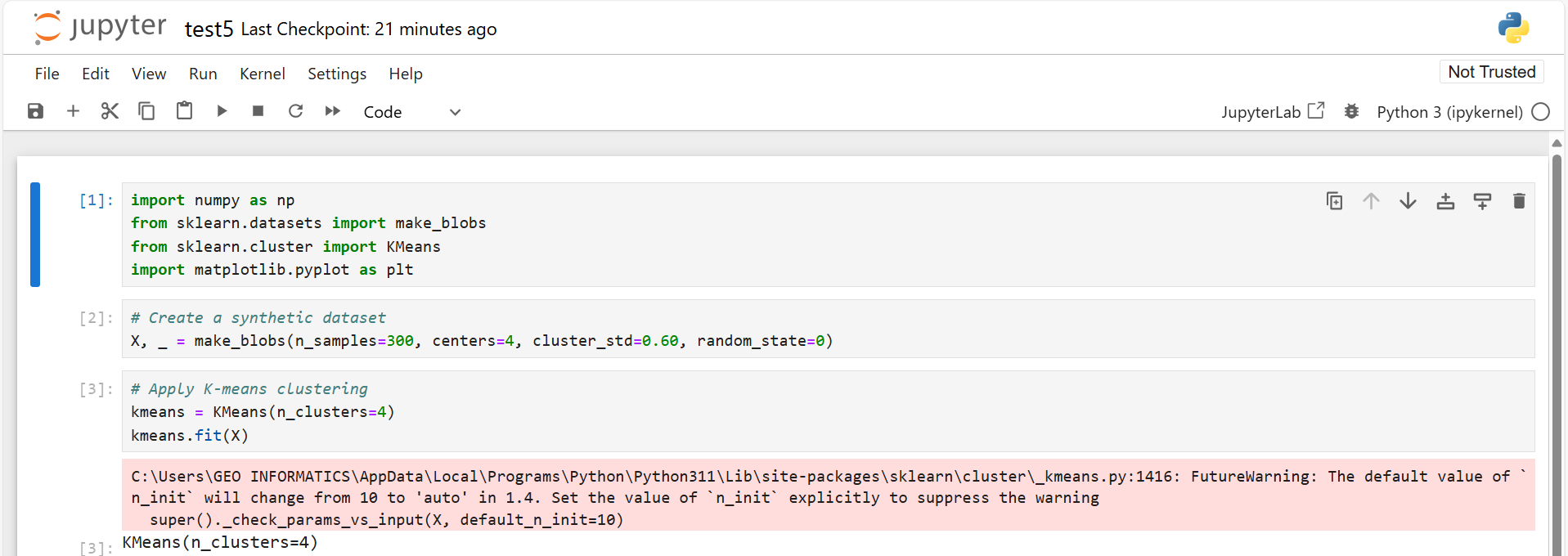
5. What is unsupervised learning? Provide two real-world applications.

Mathematical Component: Given a dataset 21,2..., write the objective function for the K-means clustering algorithm

Ans.

Unsupervised learning is a type of machine learning where the model is trained on data that does not have labeled responses. The goal is to identify patterns, structures, or relationships within the data without prior knowledge of the output values.

* 1. Customer Segmentation
  2. Anomaly Detection



6. Explain the difference between K-means and hierarchical clustering. What are the advantages and disadvantages of each?

Mathematical Component: Describe how the Euclidean distance metric is used in the K- means algorithm to update cluster centroids

Ans-:

K-means clustering is an unsupervised learning algorithm that partitions a dataset into KKK distinct, non-overlapping clusters. Each data point belongs to the cluster with the nearest mean, which serves as the cluster's centroid.

8. Discuss the difference between correlation and causation. Provide an example of a situation where correlation does not imply causation.

Mathematical Component: Given two random variables X and Y, write the formula for the Pearson correlation coefficient and describe what it measures.

Ans-: Correlation refers to a statistical measure that describes the extent to which two variables change together. It measures the strength and direction of the linear relationship between two variables.

8 Describe the difference between a probability density function (pdf) and a cumulative distribution function (cdf).

Mathematical Component: Given a continuous random variable X with pdf f(x), write the cdf F(z).

Ans\_;

The Probability Density Function (PDF), denoted as f(x)f(x)f(x), describes the probability distribution of a continuous random variable X. It specifies the likelihood of the variable taking on a specific value x.