

Feba Thomas
Roll no:50

CO1

1. Python package for numerical Calculations
2. Python package for scientific Computation
3. Python package for Data Visualizations
4. Matrix Operations using Python
5. Matrix Operations using SVD
6. What is matplotlib
7. Data visualization using matplotlib
8. What is Pandas
9. Data Handling using Pandas

1. Python package for numerical Calculations.

NumPy (Numerical Python) is the fundamental package for numerical computation in Python; it contains a powerful N-dimensional array object

2. Python package for scientific Computation

Most Python distributions include the SciPy ecosystem (open source) which includes SciPy (a SciPy library), a numerical computation package called NumPy, and multiple independent toolkits, each known as a Scikits

3. Python package for Data Visualizations

Matplotlib, seaborn, plotly, bokeh, and many more best graphing packages are available in Python for data visualization. These help in creating interactive and highly customizable plots.

4. Matrix Operations using Python

Operation on Matrix :

1. add() :- This function is used to perform element wise matrix addition.
2. subtract() :- This function is used to perform element wise matrix subtraction.
3. divide() :- This function is used to perform element wise matrix division.

5.Matrix Operations using SVD

The Singular Value Decomposition (SVD) of a matrix is a factorization of that matrix into three matrices. It has some interesting algebraic properties and conveys important geometrical and theoretical insights about linear transformations.

6.What is matplotlib

Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy.

7.Data visualization using matplotlib

Matplotlib is a low-level library of Python which is used for data visualization. It is easy to use and emulates MATLAB like graphs and visualization. This library is built on the top of NumPy arrays and consist of several plots like line chart, bar chart, histogram, etc.

8.What is Pandas

pandas is a Python package providing fast, flexible, and expressive data structures designed to make working with “relational” or “labeled” data both easy and intuitive. It aims to be the fundamental high-level building block for doing practical, real-world data analysis in Python.

9.Data Handling using Pandas

It is a package useful for data analysis and manipulation. Pandas provide an easy way to create, manipulate and wrangle the data. Pandas provide powerful and easy-to-use data structures, as well as the means to quickly perform operations on these structures.

CO2

1.What is Regression

Regression is a technique for investigating the relationship between independent variables or features and a dependent variable or outcome. It's used as a method for predictive modelling in machine learning, in which an algorithm is used to predict continuous outcomes.

2.What is Classification

In machine learning, classification is a supervised learning concept which basically categorizes a set of data into classes. The most common classification problems are – speech recognition, face detection, handwriting recognition, document classification, etc.

3.Python package for Regression

polyfit of NumPy

NumPy that stands for Numerical Python is probably the most important and efficient Python library for numerical calculations involving arrays. In addition to several operations for numerical calculations, NumPy has also a module that can perform simple linear regression and polynomial regression.

4.Python Package for Classification

scikit-learn

5.What is KNN Classifier

It is one of the simplest and widely used classification algorithms in which a new data point is classified based on similarity in the specific group of neighboring data points. This gives a competitive result.

6.How to implement KNN Classifier in Python

Build KNN classifier model. First, import the `KNeighborsClassifier` module and create KNN classifier object by passing argument number of neighbors in `KNeighborsClassifier()` function. Then, fit your model on the train set using `fit()` and perform prediction on the test set using `predict()`.

7.How to find the accuracy of the KNN Classifier

The accuracy of a classifier is given as the percentage of total correct predictions divided by the total number of instances. If the accuracy of the classifier is considered acceptable, the classifier can be used to classify future data tuples for which the class label is not known.

8.What is the Naive Bayes Algorithm, give example problems

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. For example, a fruit may be considered to be an apple if it is red, round, and about 3 inches in diameter.

9.How to find the accuracy of the Naive Bayes Algorithm

Step 1: Calculate the prior probability for given class labels.

Step 2: Find Likelihood probability with each attribute for each class.

Step 3: Put these value in Bayes Formula and calculate posterior probability. Step 4: See which class has a higher probability, given the input belongs to the higher probability class.

10.What is linear Regression

Linear Regression is a machine learning algorithm based on supervised learning. It performs a regression task. Regression models a target prediction value based on independent variables. It is mostly used for finding out the relationship between variables and forecasting.

11.What is Multiple Regression

Multiple linear regression is a statistical technique that uses multiple linear regression to model more complex relationships between two or more independent variables and one dependent variable. It is used when there are two or more x variables.

12.How to implement linear regression in Python

1. Initialize the parameters.
2. Predict the value of a dependent variable by given an independent variable.
3. Calculate the error in prediction for all data points.
4. Calculate partial derivative w.r.t a_0 and a_1 .
5. Calculate the cost for each number and add them.

6. Update the values of a_0 and a_1 .

13. How to implement Multiple Regression in Python

Multiple Linear Regression attempts to model the relationship between two or more features and a response by fitting a linear equation to observed data. The steps to perform multiple linear Regression are almost similar to that of simple linear Regression. The Difference Lies in the evaluation. We can use it to find out which factor has the highest impact on the predicted output and how different variables relate to each other.

14. How to evaluate the performance of linear and multiple regression in python

There are 3 main metrics for model evaluation in regression:

1. R Square/Adjusted R Square.
2. Mean Square Error(MSE)/Root Mean Square Error(RMSE)
3. Mean Absolute Error(MAE)

