Semester :VI_____ Subject :_____DAV_____ Academic Year: 2023 - 2024

SciPy:

Scipy is a Python library useful for solving many mathematical equations and algorithms. It is designed on the top of Numpy library that gives more extension of finding scientific mathematical formulae like Matrix Rank, Inverse, polynomial equations, LU Decomposition, etc. Using its high-level functions will significantly reduce the complexity of the code and helps better in analyzing the data.

SciPy is an interactive Python session used as a data-processing library that is made to compete with its rivalries such as MATLAB, Octave, R-Lab, etc. It has many user-friendly, efficient, and easy-to-use functions that help to solve problems like numerical integration, interpolation, optimization, linear algebra, and statistics. The benefit of using the SciPy library in Python while making ML models is that it makes a strong programming language available for developing fewer complex programs and applications

Installation of SciPy

pip install scipy

How does Data Analysis work with SciPy?

Data Preparation

- Import the necessary libraries: import numpy as np and import scipy as sp.
- Load or generate your dataset using NumPy or pandas.

Exploratory Data Analysis (EDA)

- Use descriptive statistics from SciPy's stats module to gain insights into the dataset.
- Calculate measures such as mean, median, standard deviation, skewness, kurtosis, etc.

Example:

```
from scipy import stats
data = np.array([1, 2, 3, 4, 5, 6, 7, 8, 9])
# Calculate mean and standard deviation
mean_val = np.mean(data)
std_dev = np.std(data)
# Perform basic statistical tests
t_stat, p_value = stats.ttest_1samp(data, popmean=5)
print("t_stat:" , t_stat)
print("p_value:", p_value)
```

Output

```
t_stat: 0.0
p_value: 1.0
```

Statistical Hypothesis Testing

Use SciPy's stats module for various hypothesis tests such as t-tests, chi-square tests, ANOVA, etc.

Regression Analysis

Utilize the linregress function for linear regression analysis.

Pros of using SciPy:

Semester :VI_____ Subject :_____DAV_____ Academic Year: 2023 - 2024

- Visualizing and manipulating data with high level commands and classes.
- Python sessions that are both robust and interactive.
- For parallel programming, there are classes and web and database procedures.

Cons of using SciPy:

- SciPy does not provide any plotting function because its focus is on numerical objects and algorithms.