Machine Learning Diploma

Session5: Matplotlib & Seaborn

AMIT

Agenda

➔ Matplotlib
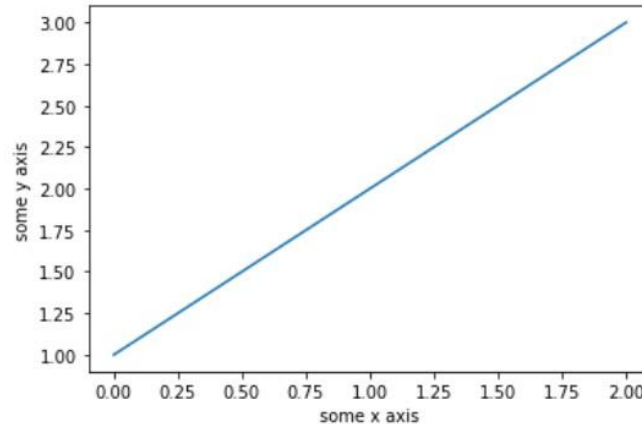➔ Seaborn
➔ Practice Project

AMIT

1. Matplotlib

Matplotlib:

➔ Matplotlib can be used for creating plots and charts. The library is generally used as follows:
- Call a plotting function with some data (e.g. .plot()).
- Call many functions to setup the properties of the plot (e.g. labels and colors).
- Make the plot visible (e.g. .show()).

AMIT

Matplotlib: Line plot

➔ To create a simple line plot from one dimensional data.

```python
import matplotlib.pyplot as plt
import numpy
myarray = numpy.array([1, 2, 3])
plt.plot(myarray)
plt.xlabel('some x axis')
plt.ylabel('some y axis')
plt.show()
```

Matplotlib: Scatter plot

➔ To create a scatter plot from two dimensional data.

```python
In [2]:  import matplotlib.pyplot as plt
         import numpy
         x = numpy.array([1, 2, 3])
         y = numpy.array([2, 4, 6])
         plt.scatter(x,y)
         plt.xlabel('some x axis')
         plt.ylabel('some y axis')
         plt.show()
```
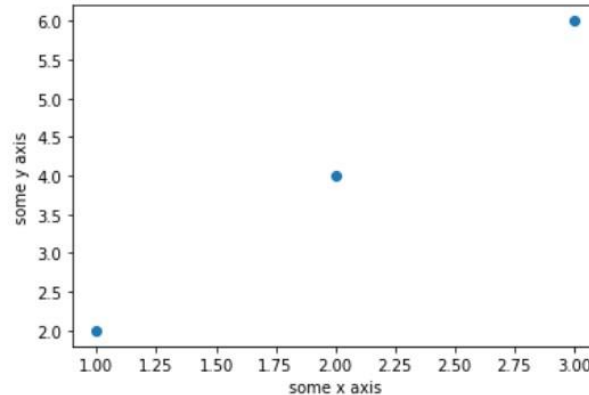
2. seaborn

Seaborn vs Matplotlib?

➜ <u>Matplotlib:</u> It is a Python library used for plotting graphs with the help of other libraries like Numpy and Pandas. It is a powerful tool for visualizing data in Python. It is used for creating statical interferences and plotting 2D graphs of arrays. It was first introduced by John D. Hunter in 2002. It uses Pyplot for providing MATLAB like interface free and open-source. It is capable of dealing with various operating systems and their graphical backends.

8 ·

Seaborn vs Matplotlib?

➔ <u>Seaborn</u>**:** It is also a Python library used for plotting graphs with the help of Matplotlib, Pandas, and Numpy. It is built on the roof of Matplotlib and is considered as a superset of the Matplotlib library. It helps in visualizing univariate and bivariate data. It uses beautiful themes for decorating Matplotlib graphics. It acts as an important tool in picturing Linear Regression Models. It serves in making graphs of statical Time-Series data. It eliminates the overlapping of graphs and also aids in their beautification.

AMIT

Seaborn vs Matplotlib?

Features	Matplotlib	Seaborn
Functionality	It is utilized for making basic graphs. Datasets are visualised with the help of bargraphs, histograms, piecharts, scatter plots, lines and so on.	Seaborn contains a number of patterns and plots for data visualization. It uses fascinating themes. It helps in compiling whole data into a single plot. It also provides distribution of data.
Syntax	It uses comparatively complex and lengthy syntax. Example: Syntax for bargraph- matplotlib.pyplot.bar(x_axis, y_axis).	It uses comparatively simple syntax which is easier to learn and understand. Example: Syntax for bargraph- seaborn.barplot(x_axis, y_axis).

3

AMIT

Seaborn vs Matplotlib?

Dealing Multiple Figures	We can open and use multiple figures simultaneously. However, they are closed distinctly. Syntax to close one figure at a time: matplotlib.pyplot.close(). Syntax to close all the figures: matplotlib.pyplot.close("all")	Seaborn sets time for the creation of each figure. However, it may lead to (OOM) out of memory issues
Visualization	Matplotlib is well connected with Numpy and Pandas and acts as a graphics package for data visualization in python. Pyplot provides similar features and syntax as in MATLAB. Therefore, MATLAB users can easily study it.	Seaborn is more comfortable in handling Pandas data frames. It uses basic sets of methods to provide beautiful graphics in python.

AMIT

Seaborn vs Matplotlib?

Pliability	Matplotlib is a highly customized and robust	Seaborn avoids overlapping of plots with the help of its default themes
Data Frames and Arrays	Matplotlib works efficiently with data frames and arrays. It treats figures and axes as objects. It contains various stateful APIs for plotting. Therefore plot() like methods can work without parameters.	Seaborn is much more functional and organized than Matplotlib and treats the whole dataset as a single unit. Seaborn is not so stateful and therefore, parameters are required while calling methods like plot()
Use Cases	Matplotlib plots various graphs using Pandas and Numpy	Seaborn is the extended version of Matplotlib which uses Matplotlib along with Numpy and Pandas for plotting graphs

AMIT

4. Practical Projects

Project: Titanic Dataset

➔ Load the dataset. And investigate its content:
- ○ What is the dimensions of this data?
- ○ How many features?
- ○ What is the output data type? Qualitative or quantitative?
- ○ Does it have nulls?
- ○ Does it have duplicate values?
- ○ Does it have outliers?
- ○ Try to make a full EDA with visualizations to investigate the data.

AMIT

Any Questions?

AMIT