Software Workshop-II

1. Numpy is very useful for fundamental scientific computations in Machine Learning. It is particularly useful for linear algebra, Fourier transform, and random number capabilities.
2. Skikit-learn is one of the most popular ML libraries for classical ML algorithms. Scikit-learn supports most of the supervised and unsupervised learning algorithms. Scikit-learn can also be used for data-mining and data-analysis, which makes it a great tool who is starting out with ML.
3. Pandas is a popular Python library for data analysis. It is not directly related to Machine Learning. It provides high-level data structures and wide variety tools for data analysis. It provides many inbuilt methods for groping, combining and filtering data.
4. matplotlib.pyplot ;-Provides a MATLAB-like plotting framework. pylab combines pyplot with numpy into a single namespace. This is convenient for interactive work, but for programming it is recommended that the namespaces be kept separate
5. Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.
6. from sklearn.model\_selection import train\_test\_split In scikit-learn a random split into training and test sets can be quickly computed with the train\_test\_split helper function.
7. from sklearn.linear\_model import LinearRegression Ordinary least squares Linear Regression. LinearRegression fits a linear model with coefficients w = (w1, …, wp) to minimize the residual sum of squares between the observed targets in the dataset, and the targets predicted by the linear approximation.
8. describe() :- Used to see the statistical details of the dataset
9. plot() :- UDE TO To see the statistical details of the dataset
10. show() :- used to show our plotted data
11. train\_test\_split() :- used to split the dataset into 2 parts(i.e. train & test)
12. plt.tight\_layout() :- Automatically adjust subplot parameters to give specified padding.
13. displot() :- Flexibly plot a univariate distribution of observations. This function combines the matplotlib hist function (with automatic calculation of a good default bin size) with the seaborn kdeplot() and rugplot() functions. It can also fit scipy.stats distributions and plot the estimated PDF over the data.
14. intercept\_ :- used to To retrieve the intercept
15. coef\_ :- used to retrieve the slope
16. predict() :- used to make predictions on the test data
17. matplotlib.pyplot.xlabel() :- Sets the label for the x-axis.
18. matplotlib.pyplot.ylabel() :- Sets the label for the y-axis.
19. dataframe plot() function it is quite easy to create decent looking plots with your dataframe
20. matplotlib.pyplot.grid() :- uesd to configure the grid lines
21. matplotlib.pyplot.scatter() :- used to scatter plot of y vs x with varying marker size and/or color.
22. NumPy sqrt () :- used to get the square root of the matrix elements.