# Pre-Requisites

| | |
|---|---|
| **DATE** | 1 NOV 2022 |
| **TEAM ID** | PNT2022TMID01032 |
| **PROJECT NAME** | WEB PHISHING DETECTION |
| **MARKS** | 2marks |

## Jupyter Lab:

Jupyter Lab is the latest web-based interactive development environment for notebooks, code, and data. Its flexible interface allows users to configure and arrange workflows in data science, scientific computing, computational journalism, and machine learning. A modular design invites extensions to expand and enrich functionality.

## sk- learn:

SK-learn is an open-source Python library that implements a range of machine learning, pre-processing, cross-validation, and visualization algorithms using a unified interface.

Simple and efficient tools for data mining and data analysis. It features various classification, regression and clustering algorithms including support vector machines, random forests, gradient boosting, k-means, etc.

## Numpy:

**Python NumPy** is a general-purpose array processing package which provides tools for handling the n-dimensional arrays. It provides various computing tools such as comprehensive mathematical functions, linear algebra routines.

NumPy provides both the flexibility of Python and the speed of well-optimized compiled C code. It's easy to use syntax makes it highly accessible and productive for programmers from any background.

NumPy is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays. It is the fundamental package for scientific computing with Python. It is open-source software.

# Pandas:

Pandas is an open-source library that is built on top of NumPy library. It is a Python package that offers various data structures and operations for manipulating numerical data and time series. It is mainly popular for importing and analyzing data much easier. Pandas is fast and it has high-performance & productivity for users.

Pandas is mainly used for data analysis and associated manipulation of tabular data in DataFrames. Pandas allows importing data from various file formats such as comma-separated values, JSON, Parquet, SQL database tables or queries, and Microsoft Excel.

# Matplotlib:

**Matplotlib** is easy to use and an amazing visualizing library in Python. It is builton NumPy arrays and designed to work with the broader SciPy stack and consists of several plots like line, bar, scatter, histogram, etc.

In this article, we will learn about ***Python plotting with Matplotlib*** from basics to advance with the help of a huge dataset containing information about different types of plots and their customizations.

If you are using anaconda navigator, follow below steps to download required packages:

1. Open anaconda prompt.
2. Type **"pip install numpy"** and click enter.
3. Type **"pip install pandas"** and click enter.
4. Type **"pip install matplotlib"** and click enter.
5. Type **"pip install scikit-learn"** and click enter.
6. Type **"pip install Flask"** and click enter.

# Flask: Web framework used for building Web applications.

There are many modules or frameworks which allow building your webpage using python like a bottle, Django, Flask, etc. But the real popular ones are Flask and Django. Django is easy to use as compared to Flask but Flask provides you with To understand what Flask is you have to understand a few general terms.

1. **WSGI** Web Server Gateway Interface (WSGI) has been adopted as a standard for Python web application development. WSGI is a specification for a universal interface between the web server and the web applications.
2. **Westernize** It is a WSGI toolkit, which implements requests, response objects, and other utility functions. This enables building a web framework on top of it. The Flask framework uses Westernize as one of its bases.
3. **Jinja2** jinja2 is a popular templating engine for Python. A web templating system combines a template with a certain data source to render dynamic web

   Pages