**5.3 SOFTWARE REQUIREMENTS**

**Python programming language:** The Python programming language is widely favoured for sentiment analysis due to its simplicity, ease of reading, and extensive library support. Python boasts a vast ecosystem of specialized libraries and tools for data analysis and natural language processing (NLP). Its NLP libraries provide robust functionality for tasks like text processing and feature extraction, which are crucial for sentiment analysis. Furthermore, Python seamlessly integrates with machine learning and deep learning frameworks like scikit-learn and TensorFlow, enabling the creation of highly accurate sentiment analysis models.

**TensorFlow:** It is an end-to-end open-source platform for machine learning. It is useful in developing and training ML models. It provides powerful deep learning capabilities, including pre-trained models that achieve high accuracy. It is scalable, allowing efficient processing of large amounts of data, and integrates well with Python.

**Keras framework:** It is a neural network library that runs on top of TensorFlow. It acts as a user-friendly python interface for artificial neural network architecture. With its wide range of pre-defined layers and model architectures, Keras provides convenient building blocks for analysing text data in sentiment analysis. Additionally, its integration with TensorFlow enables users to leverage the powerful computational capabilities and optimizations of TensorFlow, further enhancing the effectiveness of sentiment analysis tasks.

**Scikit-learn:** It is a popular ML library for the Python PL that focuses on a wide range of mathematical and statistical tools and general-purpose algorithms that form the basis of many ML technologies. It is built on top of Matplotlib which is a popular data visualization library (provides a wide range of functions and tools for creating visualizations, including plots, charts, and graphs) in Python. Scikit-learn leverages Matplotlib for visualizing various evaluation metrics and results, including the confusion matrix. Specifically, scikit-learn provides utility functions such as ‘confusion\_matrix()’ and ‘plot\_confusion\_matrix()’ that can directly generate a visual representation of the confusion matrix using Matplotlib.

**Jupyter Notebook:** Jupyter Notebook is an interactive web-based environment used for creating and sharing documents that contain live code, visualizations, and explanatory text. It supports multiple programming languages and allows users to execute code in a cell-by-cell fashion

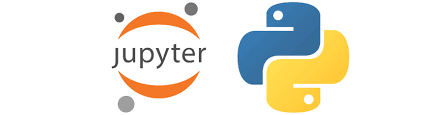


Figure 5.3.1 Jupyter and Python logo

**5.4 HARDWARE REQUIREMENTS**

A computer or server with a **multi-core CPU** and a **high-end GPU** to speed up model training and processing. It is important to consider the specific requirements of the sentiment analysis models and the libraries or frameworks that are planned to be used. Some deep learning frameworks, such as TensorFlow or PyTorch, have specific recommendations or optimizations for certain CPU or GPU architectures. Additionally, the power requirements and cooling capabilities of the CPU and GPU are to be considered to ensure stable and efficient operation.

The amount of **RAM** required to handle the size of data being processed in sentiment analysis depends on various factors, including the size of the dataset, the complexity of the models or algorithms used, and the specific requirements of the sentiment analysis tasks. It is generally recommended to have enough RAM to accommodate the size of dataset and the complexity of models. For larger-scale sentiment analysis projects, having several gigabytes to tens of gigabytes of RAM or more can be beneficial to ensure smooth and efficient processing of data.

The sufficient **storage space** required to store data and model files in sentiment analysis depends on numerous factors, including the size of the dataset, the number and size of model files, and any additional resources or metadata associated with the project. It is essential to estimate the storage requirements based on the specific characteristics of the sentiment analysis project.

**Internet connectivity** is required for accessing and processing data from external sources in sentiment analysis projects. The size of the data, the frequency of data retrieval, and any real-time processing requirements are to be considered when assessing the internet needs. A stable and reliable internet connection is essential for efficient data retrieval and processing. The recommended minimum connection speed for most sentiment analysis tasks is at least 10 Mbps. However, for large datasets or real-time streaming data, a higher connection speed may be necessary to ensure smooth and uninterrupted data access. Low latency is desirable. A latency of under 100 milliseconds is generally considered good for most sentiment analysis tasks. Sufficient bandwidth is necessary for handling the volume of data being processed. For sentiment analysis projects that involve processing large datasets or real-time streaming data, a higher bandwidth is recommended to ensure smooth and efficient data transfer.