CRC CARD:

Team Name: WE4

Class Name: numpy

Class Description:

It is a popular open-source library in Python. The foundation for many other libraries in the scientific and data science environment is NumPy, a key Python package for scientific computing.

Associated Use Cases:

1. NumPy can be used to load data from external sources like CSV files.
2. NumPy is used in the code to standardize the dataset.
3. The dataset is divided into training and testing sets using NumPy arrays.

Responsibilities:

1. NumPy is responsible for transforming the data into NumPy arrays.
2. The sophisticated array indexing and slicing features of NumPy make it possible to choose data elements or subsets.

Collaborators:

1. Pandas is often used for data loading, data manipulation, and initial data exploration.
2. collaborates with a community of open-source developers to develop, contribute, and provide support.

Class Name: pandas

Class Description:

Pandas is an open-source Python library that provides powerful and flexible data structures and data analysis tools. It is a key tool in data science and is frequently used for data preparation, modification, and analysis. Panel Data is where the name "Pandas" originates.

Associated Use Cases:

1. Pandas is used to load the diabetes dataset from a CSV file using the pd.read_csv() function
2. Tools for checking and examining the dataset are provided by Pandas.

Responsibilities:

1. Pandas are used to inspect and explore dataset.

2. Pandas are responsible for generating summary statistics for the dataset.
3. It is responsible for data exploration.

Collaborators:

1. Pandas can read data from external sources, such as CSV files, databases, and Excel spreadsheets.
2. The code uses Seaborn and Matplotlib for data display while Pandas mostly handles data manipulation.

Class Name: seaborn

Class Description:

Seaborn is an open-source Python data visualization library based on Matplotlib. It is intended especially to produce eye-catching and educational statistics visuals and data visualizations.

Associated Use Cases:

1. The "diabetes.csv" dataset's data distribution can be viewed using Seaborn.
2. To see the distribution of the "Outcome" variable, which represents diabetic and non-diabetic patients, Seaborn can be used to build a pie chart.

Responsibilities:

1. The major task for Seaborn in the code is to provide illuminating and aesthetically pleasing data visualizations.
2. The distribution of the "Outcome" variable, which represents cases with diabetes and cases without diabetes, is visualized by Seaborn.

Collaborators:

1. Matplotlib is the foundation upon which Seaborn is constructed, and it is frequently used for fine-grained plot customization.
2. Jupyter notebooks are a collaborator in interactive data analysis, albeit they are not a library.

Class Name: matplotlib

Class Description:

An open-source Python tool called Matplotlib can produce static, animated, and interactive visualizations in a variety of formats**.**

Associated Use Cases:

1. Matplotlib can be used to customize Seaborn plots created in the code.
2. Beyond what Seaborn offers, more plots can be made using Matplotlib.

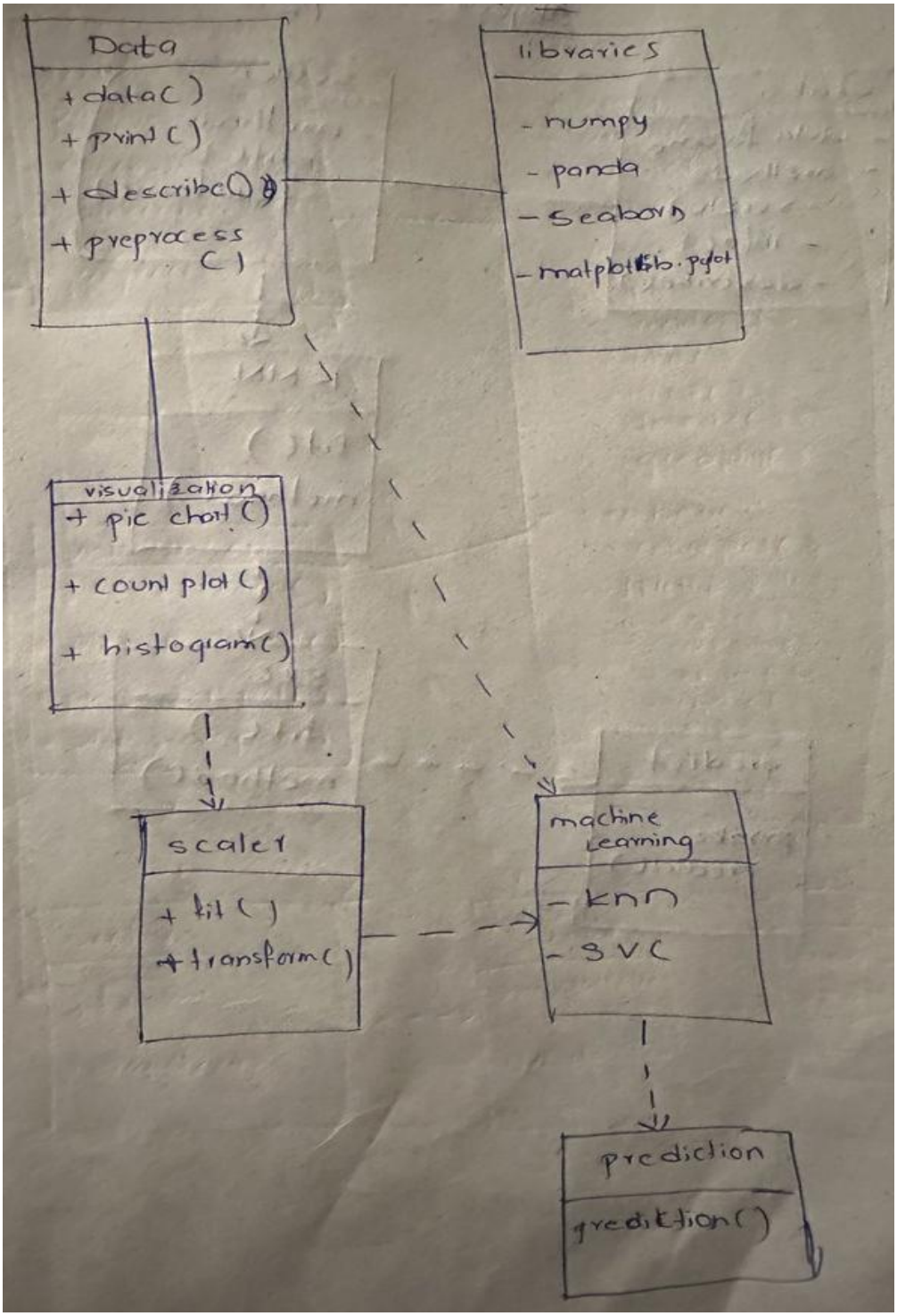
Responsibilities:

1. While static plots are the main application of Matplotlib in the code, interactive visualizations may also be made with this library.
2. Multiple Matplotlib or Seaborn plots can be combined into a single figure using Matplotlib to produce composite graphs.

Collaborators:

1. The backend of Seaborn, a higher-level data visualization library, is Matplotlib.
2. Even though it isn't explicitly demonstrated in the sample code, Matplotlib can work with Pandas Data Frames and Series.
3. The interactive environment for data processing and visualization is provided by Jupyter notebooks.

Class Diagram:



Object Diagram: