Manual for Executing or Installing My Project:

General Instructions:

1. Prerequisites:

- **Dataset Access:**
    - **Manu Siddhartha's Dataset:**
        - Access the dataset from "/content/drive/MyDrive/Colab Notebooks/Project 2 Submission – CS01080804, Afham Irfan Bin Aiman/All my codes programs/Dataset 1 v2/split_urls.csv" or download it from "https://www.kaggle.com/datasets/sid321axn/malicious-urls-dataset/download?datasetVersionNumber=1".
    - **Grega Vrbancic's Dataset:**
        - Access the dataset from "/content/drive/MyDrive/Colab Notebooks/Project 2 Submission – CS01080804, Afham Irfan Bin Aiman/All my codes programs/Dataset 2 v0/phishing-dataset-variation.csv" or download it from "Phishing-Dataset/dataset_full.csv at master · GregaVrbancic/Phishing-Dataset (github.com)".
- **Python Installation:** Make sure you have Python installed on your machine. Download and install it from https://www.python.org/downloads/ if needed.
- **Python Environment:** Choose a Python environment like Jupyter Notebook or Google Colab to run the code.

2. Code Execution:

- **Specific Instructions:**
    - **Manu Siddhartha's Dataset:** Open any .ipynb file from "/All my codes programs/Dataset 1 v2/Traditional ML" or "/All my codes programs/Dataset 1 v2/Neural Network".
    - **Grega Vrbancic's Dataset:** Open any .ipynb file from "/All my codes programs/Dataset 2 v0/models/".
- **Run Code Cells:** Run the code cells in the Jupyter Notebook sequentially (top to bottom).

- **Install Additional Libraries:** If prompted, install required libraries using `pip install <library_name>`.

3. Code Functionality:

- The code will:
    - Load the selected dataset (either Manu Siddhartha's or Grega Vrbancic's).
    - Preprocess the data (handling missing values, etc.).
    - Create features from the data using appropriate techniques.
    - Split the data into training and testing sets.
    - Train a machine learning classifier to distinguish malicious URLs.
    - Evaluate the classifier's performance using metrics like accuracy, precision, recall, and F1-score.
    - The evaluation metrics (accuracy, precision, recall, F1-score) will be displayed in the output.