SOFTWARE INSTALLATION:

Installing Python:

1. To download and install Python visit the official website of Python https://www.python.org/downloads/ and choose your version.

2. Once the download is complete, run the exe for install Python. Now click on Install Now.
3. You can see Python installing at this point.
4. When it finishes, you can see a screen that says the Setup was successful. Now click on "Close".

Installing PyCharm:

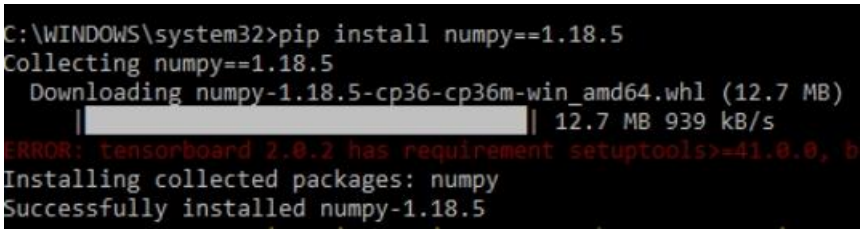
1. To download PyCharm visit the website https://www.jetbrains.com/pycharm/download/ and click the "DOWNLOAD" link under the Community Section.

2. Once the download is complete, run the exe for install PyCharm. The setup wizard should have started. Click “Next”.
3. On the next screen, Change the installation path if required. Click “Next”.
4. On the next screen, you can create a desktop shortcut if you want and click on “Next”.
5. Choose the start menu folder. Keep selected JetBrains and click on “Install”.
6. Wait for the installation to finish.
7. Once installation finished, you should receive a message screen that PyCharm is installed. If you want to go ahead and run it, click the “Run PyCharm Community Edition” box first and click “Finish”.
8. After you click on "Finish," the Following screen will appear.

9. You need to install some packages to execute your project in a proper way.

10. Open the command prompt/ anaconda prompt or terminal as administrator.

11. The prompt will get open, with specified path, type “pip install package name” which you want to install (like NumPy, pandas, sea born, scikitlearn, Matplotlib, Pyplot)



Working of the Project

The project revolves around **classifying Tor network traffic** as either **benign** or **malicious** using **Machine Learning (ML)** techniques. The workflow can be divided into several phases:

1. Data Handling

- **Input**: Network traffic data with features like Source Port, Destination Port, Protocol, Flow Duration, IAT, etc.
- **Preprocessing**: Cleaning the data by:
    - Handling missing values
    - Converting categorical data to numerical
    - Selecting relevant features

2. Model Building

Five ML models are implemented:

- **Decision Tree**
- **Logistic Regression**
- **XGBoost**
- **Random Forest**
- **Convolutional Neural Network (CNN)**

Each model learns patterns in the encrypted traffic to classify whether it belongs to **Tor or Non-Tor** traffic.

3. Model Evaluation

- Models are evaluated based on:
    - **Accuracy**
    - **Computational efficiency**
- Users can compare performance and choose the best classifier.

4. Prediction System

- Users input data (either real-time or from file).
- System predicts whether the traffic is **Tor or Non-Tor** based on trained models.

⚙️ Execution of the Project

Here's how you can run or deploy this project step-by-step:

🔸 System Requirements

Hardware:

- OS: Windows 7 or above
- RAM: 8GB minimum
- Processor: Intel i3+ or Ryzen
- Disk: 500GB HDD/SSD

Software:

- Python 3.6+
- IDE: VSCode or PyCharm
- Framework: Flask for web interface

🔸 Environment Setup

1. **Install Python**: From python.org
2. **Install dependencies** using pip:

bash

CopyEdit

pip install numpy pandas seaborn scikit-learn matplotlib flask

3. **IDE setup**: Use PyCharm or VSCode for development.

🔸 Code Modules and Functions

- **Data Preprocessing**: Cleaning and feature engineering
- **Model Training**: Train and save the ML models
- **Web UI (Flask)**:
    - Upload/View Data
    - Train Models
    - Provide Inputs for Prediction
    - Show Prediction Output

- **Running the Application**

1. Start the Flask server:

bash

CopyEdit

python app.py

2. Open browser
3. Use UI to:
    - View data
    - Train models
    - Input test traffic
    - Get classification result (Tor / Non-Tor)