

Setting Up the Environment

1. Prerequisites

Before setting up the environment, ensure you have the following installed:

- **Python 3.8+** → [Download Here](#)
- **pip (Python package manager)** (comes with Python)
- **virtualenv** (optional, for isolated environments)

To check if Python and pip are installed, run:

```
python --version  
pip --version
```

2. Create a Virtual Environment (Recommended)

Setting up a virtual environment prevents dependency conflicts.

For Windows:

Open **Command Prompt (cmd)** and run:

```
python -m venv myenv  
myenv\Scripts\activate
```

For macOS/Linux:

Open **Terminal** and run:

```
python3 -m venv myenv  
source myenv/bin/activate
```

After activation, you should see **(myenv)** at the beginning of your command line.

3. Install Required Dependencies

Once inside the virtual environment, install the required packages:

```
pip install -r requirements.txt
```

4. Verify Installation

After installation, you can verify everything is set up correctly:

```
python -c "import numpy; print(numpy.__version__)"
```

or

```
python script.py # Run the main script
```

5. Deactivate Virtual Environment (Optional)

When you're done, exit the virtual environment using:

```
deactivate
```

Alternative: Using Conda (If Preferred)

If you're using **Anaconda**, set up the environment like this:

```
conda create --name myenv python=3.8
conda activate myenv
conda install --file requirements.txt
```

Following these steps ensures a smooth setup for your project.

Run the Notebook and Reproduce Results

To run the Jupyter Notebook and reproduce results, follow these steps:

Activate the virtual environment (if not already activated):

```
source myenv/bin/activate # macOS/Linux
```

or

```
myenv\Scripts\activate # Window
```

Install Jupyter Notebook (if not installed):

```
pip install notebook
```

Launch Jupyter Notebook:

```
jupyter notebook
```

Open the provided **.ipynb** file and run all cells to reproduce results.

Alternatively, you may also open Google Colab to run the notebook in the jupyter notebook.

Key Insights and Findings

Here are some key insights and findings from our model and solutions:

Model Efficiency

After training and testing multiple models, including Random Forest, Decision Tree, and Multi-Layer Perceptron, we found that **Random Forest** performed the best in predicting both target groups: “Is Domestic Ultimate” and “Is Global Ultimate”.

The model achieved an **87% accuracy** in predicting the “Is Domestic Ultimate” group and a **91% accuracy** for “Is Global Ultimate”, demonstrating its strong predictive capabilities.

Feature Importance

The most influential features varied for each target group. When predicting “Is Domestic Ultimate”, “Industry” type was the most important factor. In contrast, for “Is Global Ultimate”, the **total number of global employees** (Employees (Global Ultimate Total)) played the most significant role.