Disease Prediction System - User Manual Guide

▼ How to Set Up After Downloading the ZIP File:

1. **Extract the ZIP file** to any preferred location on your computer.
2. Open the extracted folder.
3. You will see files like `main.py`, `train_model.py`, `model.pkl`, `dataset.csv`, etc.
4. Ensure Python is installed on your system (Python 3.8 or above).

? Project Requirements:

Before running the project, install the required Python libraries.

Run the following command in the terminal or command prompt:

```
pip install pandas numpy scikit-learn joblib
```

These are the main libraries used:

- **pandas** – for data handling
- **numpy** – for numerical operations
- **scikit-learn** – for machine learning
- **tkinter** – built-in for GUI in Python
- **joblib** or **pickle** – for saving/loading model

🚀 How to Run the Project:

1. Open terminal or command prompt in the extracted project folder.
2. Run the GUI with the following command:

```
python gui.py
```

3. The application window will open.
4. Select your symptoms from dropdowns and click **Predict Disease**.
5. The system will display:
   - Predicted disease name
   - Confidence score

- Disease description
- Precautions to follow

🔁 How to Retrain the Model:

If you want to add more diseases or symptoms:

1. Open `dataset.csv` and add the new data.
2. Save the CSV file.
3. Run the following command to retrain the model:

```
python train_model.py
```

4. A new `model.pkl` will be generated and used in the GUI automatically.

✅ If your GUI has a **"Train Model"** button, you can retrain from the GUI directly.

📂 Files and Their Purpose:

- `gui.py` – Runs the GUI interface.
- `train_model.py` – Trains the ML model using the dataset.
- `model.pkl` – Pretrained model used for prediction.
- `dataset.csv` – Main dataset with symptoms and diseases.
- `symptom_description.csv` – Description of each symptom.
- `precaution.csv` – Precautions for each disease.
- `README.md` – Overview and quick instructions.
- `requirements.txt` – Python dependencies.

📋 Validation Rules:

- Select at least 3–5 symptoms before predicting.
- Do not leave fields empty – it will show a warning.
- Ensure CSVs do not have missing or invalid data.
- Make sure the `model.pkl` file exists for predictions.

🛠️ Troubleshooting Tips:

- **GUI not opening?** Make sure Python is installed and use: `python main.py`
- **ModuleNotFoundError?** Run: `pip install pandas numpy scikit-learn`