Github Repository Setup Guide

This guide will help you set up and run the "Foundational_project_stock-Market" repository for NIFTY 50 long-term predictions.

Prerequisites

Before starting, make sure you have the following installed:

- Python 3.7 or higher
- Git
- Jupyter Notebook or JupyterLab
- pip (Python package manager)
- numpy>=1.24.4
- numba==0.57.1

Step 1: Clone the Repository

git clone https://github.com/Priyesh122/Foundational_project_stock-Market.git
cd Foundational_project_stock-Market

Step 2: Create and Activate a Virtual Environment (Recommended)

For Windows:

python -m venv venv
venv\Scripts\activate

For macOS/Linux:

python -m venv venv
source venv/bin/activate

Step 3: Install Required Dependencies

Install all the necessary packages listed in the code:

pip install pandas numpy matplotlib seaborn plotly yfinance scikit-learn shap nltk wandb streamlit jupyter

Step 4: NLTK Setup

Run Python and download the required NLTK data:

```
import nltk
nltk.download('vader_lexicon')
```

Step 5: Wandb Setup (Optional)

If you plan to use Weights & Biases for experiment tracking:

1. Sign up for a free account at wandb.ai
2. Run wandb login in your terminal and enter your API key

Step 6: Running the Jupyter Notebook

1. Open the Jupyter Notebook:

```
jupyter notebook
```

2. Navigate to and open FP_SMP_LTv3.ipynb
3. Run all cells in the notebook to:
    - Download historical NIFTY 50 data
    - Preprocess the data
    - Train the prediction models
    - Generate visualizations and insights
    - Save the trained model

Step 7: Running the Streamlit App

After running the notebook, you can launch the interactive Streamlit app:

```
streamlit run streamlit_app.py
```

This will start a local web server, typically at http://localhost:8501, where you can interact with the predictions through a user-friendly interface.

IMPORTANT NOTE: Please keep **ALL** the files in the same directory and do **NOT** change the title (even the logo file)

File Descriptions

- **FP_SMP_LTv3.ipynb**: The main Jupyter notebook containing the data processing, model training, and evaluation
- **long_term_investor_model.py**: Python module for the long-term investor prediction model
- **long_term_modelv3.pkl**: Pickled (saved) trained model
- **streamlit_app.py**: Streamlit application for interactive visualization of predictions
- **README.md**: Project documentation
- **Foundational-Project.pptx**: Presentation slides for the project
- **nifty50_logo.png**: Logo image used in the application

Troubleshooting

SSL Certificate Issues

If you encounter SSL certificate validation errors (especially on macOS), the code includes a workaround:

```
import ssl
ssl._create_default_https_context = ssl._create_unverified_context
```

This is already included in the imports, but if you're running parts of the code separately, you may need to include this.

yfinance Data Download Issues

Important: The yfinance library may throw errors during stock market opening hours due to high server load or rate limiting. If possible, try running the data download portions of the notebook during off-market hours for more reliable results.

If you encounter errors like "HTTPError: HTTP Error 429: Too Many Requests" or other connection issues, wait a few minutes and try again, or run the code during market off-hours.

Package Installation Issues

If you encounter issues installing any packages, try installing them individually:

```
pip install problematic-package
```

Or specify a version if needed:

pip install problematic-package==specific-version

Notes

- The model uses historical NIFTY 50 data for predictions
- Make sure you have a stable internet connection when running the notebook as it downloads data from Yahoo Finance
- The SSL certificate workaround is included to handle potential issues with data downloads