Mini Project: Predicting Stock Price Movements Using Convolutional Neural Networks

Start Assignment

- Due Mar 27 by 6:30pm
- Points 100
- Submitting a file upload
- Available after Feb 20 at 6:30pm

Project Overview

This project is a simplified implementation of the paper *"(Re-)Imag(in)ing Price Trends"* by Jingwen Jiang, Bryan Kelly, and Dacheng Xiu. The objective is to train a Convolutional Neural Network (CNN) to predict binary stock price movements using image-based market data. Instead of manually identifying price patterns, the model will automatically learn and extract features that are predictive of future price movements.

Students can work in groups of up to three people. Please sign up for their groups on Canvas (Canvas > People > Groups > Mini Project) before **6:30 PM, March 6, 2025**.

You are provided with:

- Image data: Stock price charts formatted as images.
- Labels: Binary labels indicating whether the stock price increased (1) or decreased (0) in the following period.
- Starting Code: You will be given starter code that demonstrates how to read the image data and implement a simple CNN for prediction. The CNN model serves as a naive baseline.

You can access the dataset and starting code at the following OneDrive folder [**Link to Ondrive (https://stevens0-my.sharepoint.com/:f:/g/personal/zyang99_stevens_edu/EvwZTTKgFuFFkvImOSYrY-0BOPj74T_IbMkfMkKNOzuz4Q?e=gSY43g)**].

Project Tasks

1. Data sampling:
   - Split the dataset into training, validation, and test sets.
   - You can apply time-series split or rolling-/expanding-window split.
1. Build a CNN Model
   - Implement a CNN architecture suitable for stock price movement classification.
   - Train the model using the provided dataset to predict the five-day binary return trend.
   - Experiment with different hyperparameters (e.g., number of convolutional layers, kernel size, activation functions, dropout, batch size).

2. Evaluate the Model
    - Use accuracy, precision, recall, and F1-score to assess model performance.
    - Visualize training and validation loss curves.
    - Analyze misclassified examples and discuss potential improvements.
3. Report Writing
    - Describe your methodology and implementation details.
    - Discuss model performance and insights from the results.
    - Provide recommendations for potential improvements or future extensions.

Deliverables

Each group must submit:

1. A Written Report (PDF format), including but not limited to the following:
    - Introduction: Briefly introduce the project, motivation, and objectives.
    - Data Description: Explain the dataset and preprocessing steps.
    - Methodology: Describe the CNN architecture and training process.
    - Results: Present model evaluation metrics and visualizations.
    - Conclusion: Summarize key findings and propose future improvements.
2. Code Implementation (Jupyter Notebook or Python script)
    - Clearly commented and structured.
    - Include necessary instructions for running the code.

Grading Criteria

- Report Quality (80%): Clarity, structure, completeness, analysis, and discussion.
- Code Readability (20%): Well-documented and organized code.
- All group members will receive the same grade.

Reference

Jiang, J., Kelly, B., & Xiu, D. (2023). (Re-) Imag (in) ing price trends. *The Journal of Finance, 78(6),* 3193-3249.