
Stock Prediction with Price Trend Graph

Project 1, Spring 2025

Deadline: 2025.04.23 22:00 (Not 23:59)

1 Data Description

In this project, we provide you data of A-share stock price trend in image format.

1.1 Image Dataset

The image dataset includes about 90,000 randomly sampled images over stock trade records in A-share market. Specifically, we include data from 2010-2012 in training set and data from 2013 in test set. All images have sizes of 180 * 96.

A typical price trend image looks like Figure 1. A stock's daily open, close, high and low price with trading volume are included in one graph. Such images were preprocessed according to method from (1), and the corresponding processed data is shown in Figure 2. Each image includes price and volume data of 64 days, and the starting price is set to 1 for all the graphs.



Note: OHLC chart for Tesla stock with 20-day moving average price line and daily volume bars. Daily data from January 1, 2020 to August 18, 2020.

Figure 1: A Tesla stock price trend image

The structure for image dataset is organized as follow:

- train: Folder of training set images, ranging from 2010 to 2012. The folder included two subdirectories: 0 and 1, which correspond to images with negative 5-day ahead return and positive 5-day ahead return respectively.
- test: Folder of test set images, including data from 2013. The folder included two subdirectories: 0 and 1, which correspond to images with negative 5-day ahead return and positive 5-day ahead return respectively.



Figure 2: An A-Share stock's processed price trend image

2 Task

Given the image dataset, please create a CNN-based network for classifying the 5-day ahead return of each image into positive class or negative class. The result should be the classification accuracy of your model over test dataset. Please also report the structure and hyperparameters used in your network, and analyze how they affect the accuracy.

Hint: You can use “ImagerFolder” from PyTorch to directly load the image dataset, we provide a starting code `main.py` for you to load the images from train and test datasets. The model structures used in original paper (1) are provided in Figure 3. You can inherit their CNN network, or create your own CNN network. There is no requirement for creating a validation set, it is up to you to decide how to split your training set.

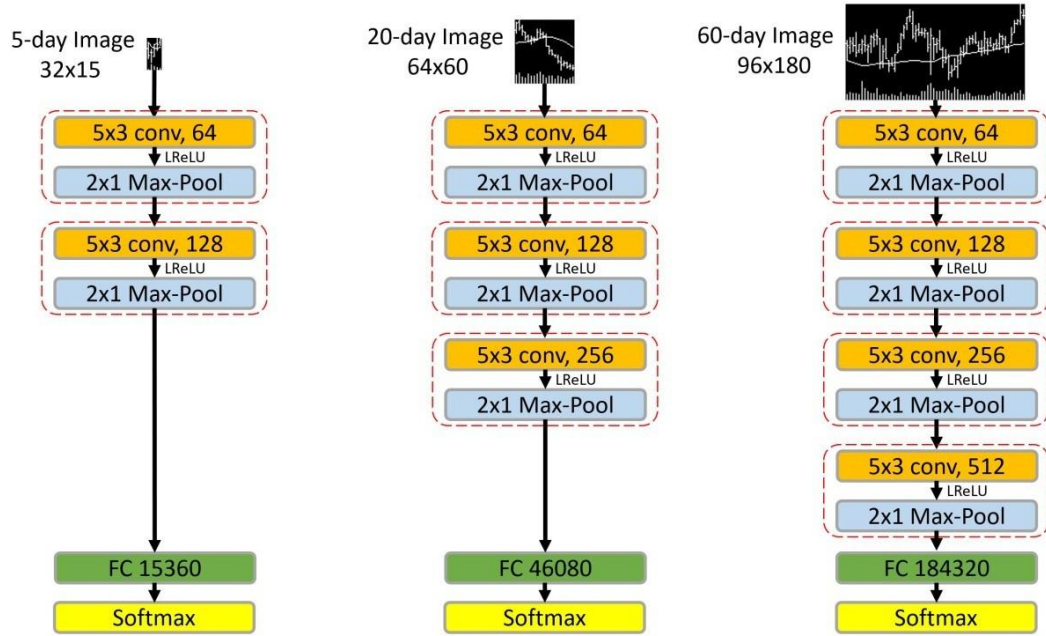


Figure 3: Networks from original paper

3 Requirement

You need to submit all codes and a short report (pdf format) with the following requirements:

- Completing and successfully running baseline networks for the task .
- Including out-of-sample accuracy for the task in your report. You can also support it with tables like confusion matrix and do performance analysis with metrics like precision and recall rate. Do analysis on how different model structures, hyperparameters affect the result.
- Exploring the pros and cons from using CNN over images to predict stock returns.

In case you have no idea how good your result is, here is the accuracy from the original paper:

Out-of-Sample Classification Accuracy

Table 2: Out-of-Sample Classification Accuracy

Image size	Return horizon			
	20-day		60-day	
	Acc.	Corr.	Acc.	Corr.
5-day	52.5%	3.1%	53.6%	2.3%
20-day	53.3%	3.4%	53.2%	2.4%
60-day	53.6%	2.9%	52.9%	3.1%
MOM	52.1%	1.8%	52.1%	1.5%
STR	50.4%	1.4%	49.8%	1.2%
WSTR	51.0%	2.8%	50.4%	2.6%

Note: The table reports out-of-sample forecast performance for image-based CNN models and benchmark signals. We calculate **classification accuracy and correlation cross-sectionally each period** then report time series averages over each period in the test sample.

4 Attention

- You need to submit all codes and a report (at least two pages in PDF format).
- The report should show all improvement attempts (whether successful or not) and corresponding results in detail. We encourage you to try with more advanced CNN-based networks like ResNet in addition to tuning hyperparameters. Provide an analysis of the results.
- Plagiarism is not permitted. If open-source code is borrowed, be sure to include a description.

References

- [1] JIANG, J., KELLY, B., AND XIU, D. (re-)imag(in)ing price trends. *The Journal of Finance* 78, 6 (2023), 3193–3249.