**Peer review sheet**

MAFS6010Z, 2021 fall

Your name and sid: YE Xiaoyu 20814251

Group that you review: 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Confidence on your assessment (1-3) | Clarity and quality of writing (1-5) | Technical quality  (1-5) | Overall rating  (1-5) |
| Score | 3 | 3.5 | 4.5 | 4 |

**Summary:**

This project not only replicated the baseline result based on 20-day images, but also conducted a sensitivity analysis of several hyperparameters, and compared the prediction result of other time horizons (1-day, 5-day and 60-day).

**Strengths:**

Very comprehensive work, including the basic task as well as some extensions, we can see the great effort and time they spend. And especially for the “Prediction for Other Labels” (different time horizons), a strong and meaningful conclusion is drawn.

**Weaknesses:**

This project only shows the result of different train and test results, while excluding solid analysis and evaluation regarding the results. For example, what is the potential reason for the worse performance? We can see from Figure 5 that with the change of dropout rate, removal of Xavier and change of the activation function, there is a potential overfitting problem that makes the model in vain.

**Clarity and writing:**

The report is clear and easy to follow, which shows the workflow, results, and conclusion in a good manner.

However, a better structure of the 2nd part, “The CNN Model” can be used. The report puts the data processing between the model construction and the training process, which is a bit weird. Also, the visualization of the train, validation and test losses is confusing (Figure3,4). It’s not that clear that if the horizontal axis stands for iterations just as in figure 2. If not, it’s better to convert them into similar scales and granularity (figure 2 is fluctuating which figures 3&4 are smooth).

In the 1st paragraph of “2.2 Prediction for Other Labels”, the report mentions the highest accuracy is found to be that of predicting daily returns, however, it does not mention the test accuracy, which makes the result doubtful (potential overfitting again).

**Technical quality:**

Clear coding and comprehensive work. Need further tuning of the model to get better accuracy.