

Phase 2

Copy the code from the previous phase into a new file called phase2.py, and save it in the same directory.

We have the AAPL stock prices snapshots taken for the last 5 years. I would like you to do the following set of tasks. Please provide the code for all the steps.

Task 1: Prepare the repository for your phase 2.

1. Now, you have a git repo. Please reorganize the repository and move the script that you have created for phase 1 to a folder under the repository called: phase1.
2. Please save the text file for the previous phase to markdown page under phase 1 folder.
3. Create a new directory under the same repository called phase 2 and add the solution for this phase into that repository.
4. Please document all the git commands used to markdown page under phase 2 folder.

Task 2: Analysis of AAPL stock.

1. Download the AAPL.csv file stocks file.
2. Parse the date column and create a new column that has the date name like as: "mon", "tue", "wed", etc... Name that column: "day_of_week".
3. Create a new column that that is the percentage of difference between the "Open" and "Close" on the same day with the name "same_day_delta".
4. Create a new column that is either: 0 or 1. The value is set to 0 when "same_day_delta" is less than or equal to 0.0, and it's set to 1 when it's larger than that value. Name that column "same_day_strategy".
5. Create a new column that is the percentage for change between "Close" and last trading day "Close". Name that column "next_close_delta".
6. Create a new column that is either: 0 or 1. The value is set to 0 when "next_close_delta" is less than or equal to 0.1, and it's set to 1 when it's larger than that value. Name that column "next_close_strategy".
7. Find the following information per month and save it to a file: "monthly_analysis.csv":
 - a. Average close price.

- b. Average open price.
 - c. Highest close price.
 - d. Lowest open price.
 - e. Highest high and low price.
 - f. Lowest high and low price.
8. Create a Jupyter notebook and plot the both “open” and “close” prices.
 9. Plot the “same_day_delta”.
 10. Plot the distribution of “same_day_delta”.

Task 3: Machine learning. Accuracy of those models is not important.

1. Create a machine learning model that predicts: “same_day_strategy”. Please make sure to document everything you do and how you built the model and your choices in a separate markdown page “same_day_model_documentation.md”. Document the accuracy and criteria used in creating this model.
2. Create a machine learning model that predicts: “next_close_strategy”. Please make sure to document everything you do and how you built the model and your choices in a separate markdown page “same_day_model_documentation.md”. Document the accuracy and criteria used in creating this model.
3. Save those models to disk and submit them to repo.

Task 4: Optimize the inference time of the model.

1. Measure the execution time for processing the data and the inference time of the models.
2. Try to optimize the time above. Try to compare to another code.