CS7646 ML4T
Machine Learning
for Trading

≡

Q TRADER HINTS

OVERVIEW

In this project you will apply the Q-Learner you developed earlier to the trading problem. Note that there is no regression or classification learning in this project (so no use of RTLearner or LinRegLearner). The indicators define most of the "state" for your learner, an additional component of state you may use is whether or not you are currently holding a position long or short. The recommended actions are LONG, CASH, SHORT.

Overall, your tasks for this project include:

- Build a strategy learner based on your Q-Learner and previously developed indicators.
- Test/debug the strategy learner on specific symbol/time period problems

IMPLEMENT STRATEGY LEARNER

For this part of the project you should develop a learner that can learn a trading policy using your Q-Learner. You should be able to use your Q-Learner from the earlier project directly, with no changes. You will need to write code in `StrategyLearner.py` to "wrap" your Q-Learner appropriately to frame the trading problem for it. Utilize the template provided in `StrategyLearner.py` Overall the structure of your strategy learner should be arranged as below. Note that this is a suggestion, not a requirement:

For the policy learning part:

- Select several technical features, and compute their values for the training data
- Discretize the values of the features
- Instantiate a Q-learner
- For each day in the training data:
  - Compute the current state (including holding)
  - Compute the reward for the last action

- Query the learner with the current state and reward to get an action
- Implement the action the learner returned (LONG, CASH, SHORT), and update portfolio value

- Repeat the above loop multiple times until cumulative return stops improving.

For the policy testing part:

- For each day in the testing data:
  - Compute the current state
  - Query the learner with the current state to get an action
  - Implement the action the learner returned (LONG, CASH, SHORT), and update portfolio value
  - DO NOT UPDATE Q — learning must be turned off in this phase

- Return the resulting trades in a data frame

Training and testing for each situation should run in less than 30 seconds. We reserve the right to use different time periods if necessary to reduce auto grading time.