kaggle

Host   Competitions   Scripts   Jobs   Community ▾   Rupak Chakraborty   Logout

Thank you for accepting the rules.

$50,000 • 549 teams

The Winton Stock Market Challenge

Merger and 1st Submission Deadline

Tue 27 Oct 2015   Tue 26 Jan 2016 (57 days to go)

Dashboard

Home
| Data
| Make a submission

Information
| Description
| Evaluation
| Rules
| Prizes
| About Winton
| Timeline

Forum

Leaderboard

My Team

My Submissions

Leaderboard

1. Bill
2. Humberto Brandão
3. alsimak
4. Dollars4Donuts
5. Tim Hochberg
6. Rahan
7. TomM
8. Tharsis(μ,σ²)
9. Mendrika Ramarlina
10. FrMa

Forum (35 topics)

Evaluation Function
11 hours ago

domain knowledge
yesterday

Arbitrary Time Window
yesterday

Is this data valid?

Competition Details » Get the Data » Make a submission

Data Files

| File Name	Available Formats
sample_submission.csv	**.zip (7.74 mb)**
train.csv	**.zip (72.69 mb)**
test.csv	**.zip (73.73 mb)**

In this competition the challenge is to predict the return of a stock, given the history of the past few days.

We provide 5-day windows of time, days D-2, D-1, D, D+1, and D+2. You are given returns in days D-2, D-1, and part of day D, and you are asked to predict the returns in the rest of day D, and in days D+1 and D+2.

During day D, there is intraday return data, which are the returns at different points in the day. We provide 180 minutes of data, from t=1 to t=180. In the training set you are given the full 180 minutes, in the test set just the first 120 minutes are provided.

For each 5-day window, we also provide 25 features, Feature_1 to Feature_25. These may or may not be useful in your prediction.

Each row in the dataset is an arbitrary stock at an arbitrary 5 day time window.
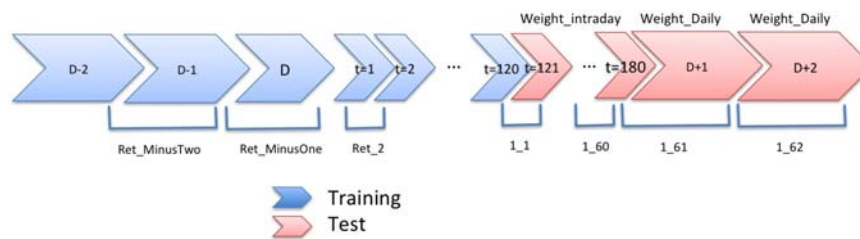
How these returns are calculated is defined by Winton, and will not to be revealed to you in this competition. The data set is designed to be representative of real data and

teams

players

entries

so should bring about a number of challenges.

File descriptions

- **train.csv** - the training set, including the columns of:

  - Feature_1 - Feature_25
  - Ret_MinusTwo, Ret_MinusOne
  - Ret_2 - Ret_120
  - Ret_121 - Ret_180: **target variables**
  - Ret_PlusOne, Ret_PlusTwo: **target variables**
  - Weight_Intraday, Weight_Daily

- **test.csv -** the test set, including the columns of:

  - Feature_1 - Feature_25
  - Ret_MinusTwo, Ret_MinusOne
  - Ret_2 - Ret_120

- **sample_submission.csv** - a sample submission file in the correct format

Data fields

- **Feature_1 to Feature_25:** different features relevant to prediction
- **Ret_MinusTwo:** this is the return from the close of trading on day D-2 to the close of trading on day D-1 (i.e. 1 day)
- **Ret_MinusOne:** this is the return from the close of trading on day D-1 to the point at which the intraday returns start on day D (approximately 1/2 day)
- **Ret_2 to Ret_120**: these are returns over approximately one minute on day D. Ret_2 is the return between t=1 and t=2.
- **Ret_121 to Ret_180**: intraday returns over approximately one minute on day D. **These are the target variables you need to predict as {id}_{1-60}.**
- **Ret_PlusOne**: this is the return from the time Ret_180 is measured on day D to the close of trading on day D+1. (approximately 1 day). **This is a target variable you need to predict as {id}_61.**
- **Ret_PlusTwo**: this is the return from the close of trading on day D+1 to the close of trading on day D+2 (i.e. 1 day) **This is a target variable you need to predict as {id}_62.**
- **Weight_Intraday**: weight used to evaluate intraday return predictions Ret 121 to 180
- **Weight_Daily**: weight used to evaluate daily return predictions (Ret_PlusOne and Ret_PlusTwo).

About   Our Team   Careers   Terms   Privacy   Contact/Support