

# ERP Prediction Contest - Rules

## 1 *Best Prediction Category*

The rules for this category are as follows.

1. The target variable for the prediction problem is “ASPFWR5” as provided in the Hull Tactical data set. This is the 5 (trading) day forward ERP of the S&P 500 adjusted for dividends.
2. This category of the contest will have two evaluation phases. Phase I will involve an out-of-sample test and Phase II will involve a real-time prediction scenario.
3. Phase I of the contest will be judged on the best R-square measure of your prediction algorithm based on some period in 2018 (as a result the data set you are given does not include 2018). The last date for submitting your algorithm for phase I will be April 7th, 11:59 P.M.
4. Phase II involves a live contest that will run from April 8th to May 8th. It will evaluate your prediction algorithms on real time S&P 500 returns. You will be able to update your code at the end of every day for following day prediction.
5. Phase II will include a live leaderboard. These live scores will again be based on R-squared performance.
6. The winner of the Best Prediction category will be based on the average of the final R-square scores from the two phases.
7. Using additional data sources is also allowed provided you source them programmatically. Your final implementation will be made public under a free software license.

### 1.1 Phase I evaluation details:

Here we provide an example for evaluation of Phase I.

- Let us assume your prediction model is  $Y_t = \beta_0 + \beta_1 X_t^1 + \beta_2 X_t^2 + \epsilon$ , where
  - $Y_t$  is 5 (trading) day forward ERP of the S&P 500 “ASPFWR5<sub>t</sub>”.
  - $X_t^1, X_t^2$  are independent variables, say  $BDIY_t$  and “VIX<sub>t</sub>” respectively.
- Let us also assume the chosen period for evaluation of your model for this phase is March 5, 2018 (Monday) - June 29, 2018 (Friday).
- First prediction of 5 (trading) day forward ERP of the S&P 500  $\hat{Y}_1$  happens on the morning of March 5, 2018 (using data for  $BDIY$  and  $VIX$  at the end of last trading day before March 5, 2018). Corresponding realized ERP  $Y_1$  will be computed in the morning (before market opens) on March 12, 2018.
- Similarly, prediction  $\hat{Y}_2$  on March 6, 2018 will be realized  $Y_2$  on March 13, 2018.
- Finally, prediction  $\hat{Y}_T$  on June 29, 2018 will be realized  $Y_T$  on July 6, 2018.
- With two time series  $(Y_1, Y_2, \dots, Y_T)$  and  $(\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_T)$ , we evaluate the performance of your model for phase I using:  $R_I^2 = 1 - \frac{\sum_{t=1}^T (Y_t - \hat{Y}_t)^2}{\sum_{t=1}^T (Y_t - \bar{Y})^2}$ , where  $\bar{Y} = \frac{1}{T} \sum_{t=1}^T Y_t$ .

## 1.2 Phase II evaluation details:

For this contest, we will have a live leaderboard updated everyday at 9:00 pm. You can update your model everyday by mid-night.

1. At 4:00 am PST every day, we will use your model for 5 (trading) day forward ERP prediction starting April 8, 2019 (Monday) until morning of May 9, 2019 (Thursday).
2. The prediction for the last day of the contest (May 9, 2019) will be realized in the morning on May 16, 2018.
3. With two new time series  $(Z_1, Z_2, \dots, Z_T)$  and  $(\hat{Z}_1, \hat{Z}_2, \dots, \hat{Z}_T)$ , we evaluate the performance of your model for phase  $I$  using:  $R_{II}^2 = 1 - \frac{\sum_{t=1}^T (Z_t - \hat{Z}_t)^2}{\sum_{t=1}^T (Z_t - \bar{Z})^2}$ , where  $\bar{Z} = \frac{1}{T} \sum_{t=1}^T Z_t$ .
4. The live leader board will start from April 16th, 9:00 pm.

## 1.3 Final Winners and Presentation:

1. All the submission will be ranked using  $\mathbf{R}^2 = \frac{R_I^2 + R_{II}^2}{2}$ .
2. Winners will deliver a presentation to explain their ideas on May 16th.