(1) Data

- Load 1-min

- Load daily

- Add fut\_cur\_gen\_ticker to 1-min

- Convert 1-min to DateTimeIndex

(2) Responses

- 1, 5, 10, 20, 40, 80 min returns (value)

- 1, 5, 10, 20, 40, 80 min up or down (class)

(3) Features

- Hour (of day)

- DayOfWeek

- Close -> Open change

- 1, 5, 10, 20, 40, 80 min prior return

- 1, 5, 10, 20, 40, 80 min cum volume

- 1, 5, 10, 20, 40, 80 min signed volume

- 1, 5, 10, 20, 40, 80 min ema, price vs ema

- price as percentile of prior 5, 10, 20, 40, 80 min range

(4) Classification results (predict up/down)

- Methodology:

- cross-validation

- assess fit

- Algos:

-kNN

-decision trees

- naïve Bayes ?

- Logistic regression

- SVM

- adaBoost?

(5) Predicting numeric values (returns, or price)

- Methodology:

- cross-validation

- assess fit

- Algos:

- regression

- tree-based regression ?