

HWK 4:

Read the documentation about PyCaret (<https://pycaret.gitbook.io/docs/>) – PyCaret is an algorithm evaluation library. For example, it compares different regression algorithms and picks the best-performing one.

Day 1:

Download for 10 consecutive hours the information about the eight currency pairs below, aggregated at the 6-minutes level – the same way you have done on the previous homework (Timestamp, MAX, MIN, MEAN, VOL, FD, and Return – use **return=0** for the first 6-minutes bucket return):

EURUSD, GBPUSD, USDCHF, USDCAD, USDHKD, USDAUS, USDNZD, and USDSGD. (This dataset will be used for training purposes.)

(Download it yourself; don't use your friend's dataset!)

Because you have 100 data points for each currency pair, you will statistically classify the VOL and the FD into 3 different classes for each currency pair:

1. High VOL and high FD (top 33 data points);
2. Medium VOL and high FD (following 34 data points); and
3. Low VOL and high FD (lower 33 data points).

Build a dictionary and replace the original values of VOL and FD with 1 (low FD and low VOL), 2 (medium FD and medium VOL), and 3 (high FD and high VOL).

Day 2:

The objective is to use the trained dataset from Day 1 to predict in real-time the next hour's return on Day 2 using VOL, FD, and MEAN (3 features). (Based on the classification of FD and VOL (from Day 1), you should automatically replace them in real-time with 1, 2, or 3.)

1. Use a regression algorithm in your prediction, like the long-jump example discussed in class. The independent variable is the VOL, FD, and MEAN matrix; the dependent variable is the hourly Return vector.
2. Use the best regression sorted by PyCaret in your prediction – it is essential to understand how to use the library.
3. The output of HWK 4 is a *.csv file with the predicted return, the actual return, and the error.
4. In HWK 5 (next homework), we will use the errors of the return estimates to improve the results gradually (i.e., to improve the next hour's real-time return estimates based last hour's real-time errors). (In HWK 5, you will also start to think about a profitable trading strategy.)

Good luck! I will schedule a section on **Thursday morning** (12/01/2022) at NYU for questions.