

Predict best which permutation settings to use

CrowdWisdomTrading ML programmer position assessment:

Hi,

The document gives a brief introduction to an assignment to check your qualification for the internship. Please complete it on your own time and submit it within 1 week.

Prerequisites:

python, google colab, Analytical thinking, ML

Problem Statement:

We have a list of trades of different simulations (each is numbered and based on different settings). Each simulation represents a different permutation of settings used to run them. These are results of a trading algorithm which decides when and how to buy or sell different assets.

We want to know what permutation will work for Wednesday 15:00 (for example) and we assume that each day of the week and hour of the day has a different profile as certain activities are based on news which are the same hours of the days. etc.

Selecting the best permutation (best by aggregated profit & loss) for a future period is needed by selecting day of the week and time of the day. You can take any feature you want, and you can use anything which will make the prediction better (you can even download Gold prices and use its value and volatility if you think this might help) so long as your prediction is the profit/loss.

The different ML method / model is upon your selection.

All work will be done on google colab

Additional reading material:

<https://qoppac.blogspot.com/2015/11/the-three-kinds-of-overfitting.html>

<https://sarit-maitra.medium.com/take-time-series-a-level-up-with-walk-forward-validation-217c33114f68>

<https://towardsdatascience.com/ai-in-finance-how-to-finally-start-to-believe-your-backtests-1-3-1613ad81ea44>

<https://www.kaggle.com/cworsnup/backtesting-cross-validation-for-timeseries>

<https://quant.stackexchange.com/questions/50033/difference-between-cross-validation-backtesting-historical-simulation-monte-c>

<https://machinelearningmastery.com/backtest-machine-learning-models-time-series-forecasting/>

<https://medium.com/cindicator/backtesting-time-series-models-weekend-of-a-data-scientist-92079cc2c540>

<https://ideas.repec.org/jmp/2018/plo493.pdf>

<https://towardsdatascience.com/time-based-cross-validation-d259b13d42b8>

<https://medium.com/@soumyachess1496/cross-validation-in-time-series-566ae4981ce4>

<https://hub.packtpub.com/cross-validation-strategies-for-time-series-forecasting-tutorial/>

The task is to do the following:

1. Build a prediction model that is based on historical data will decide which model to use per specific hour of the day for the next period of time (week? Day? Month? -

your model to decide). the prediction is for the "Profit/Loss" (the results should be a table format telling me which simulation to use on what day/hour)

	Sunday	Monday
00:00	sim 105	don't trade
01:00	sim 101	sim 104
02:00		

2. create aggregated pn/l results (with a list of trades of the predicted simulation for the OOS) of the rolling prediction (out of sample) using cross validation which is both visualised as chart and as a table of predicted vs actual (on a daily format).

date	predicted sim	actual profit/loss	aggregated profit
1, 00:00	105	55\$	55\$
1 01:00	104	-15\$	40\$
.....
100 23:00	107	12\$	56655\$

You should submit a google colab which will work with no bugs

Estimated Work Hours: 12-15 hrs

For submissions / queries, pls email to gilad@crowdwisdomtrading.com

Thanks for applying!

Gilad

CrowdWisdomTrading CEO