Reinforcement Learning For Financial Data

Stanford University

URL: https://vimeo.com/user105830642/review/376931074/b2108940fb

Introduction

Financial data like stock returns are known to be extremely noisy and difficult to predict. In the following study, we try to apply Reinforcement Learning approaches to test the efficiency of a trained trading agent on predicting future stock returns, and we assess its predictive power in a trend following strategy.

Data, Baseline and first signals

We consider the daily returns of the SP500 in the following examples (we have conducted the same study on Amazon, Google… and other stocks, obtaining similar results).

We define a first Baseline trading agent, and a first ML agent as follow:

Baseline (M1):

- If the stock is up at t, we assume is will be up at t+1 and we invest in the stock
- If the stock is down at t, we short it at t+1

First ML Agent (M2):

We design two new features to estimate the trend (or drift) of the returns of our stock: Short Moving Average and Long Moving Average.

- If the short average is larger than the long, we buy the stock
- Otherwise, we short it

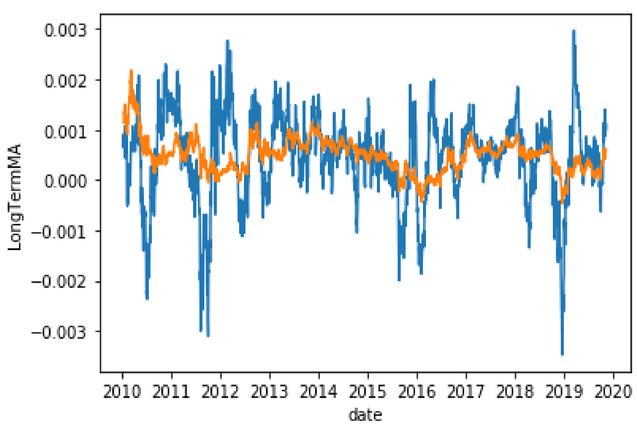

Figure 1. Long (orange) vs Short (blue) Moving Averages

Second ML Agent (M3):

We design an additional feature, the short term Volatility of the stock, and we base our trading strategy on the Linear Regression of the stock returns against the set of feature.

Interestingly, we find beta coefficient for the short and long term moving averages that are close to 1 and -1 respectively, showing that our previous intuitive approach is confirmed by a data-driven model.

We go short of the stock if our prediction is negative, and long otherwise.

We compare different Regression models (OLS, LASSO, Ridge) according to the methodology in the Testing Approach paragraph, and we keep the Ridge in our final version.

date	SPX	DailyR	ShortTermMA	LongTermMA	ShortVol
2001-03-28	1153.29	-0.024430	-0.002133	-0.000966	0.246222
2001-03-29	1147.95	-0.004630	-0.001743	-0.000986	0.240308
2001-03-30	1160.33	0.010784	-0.002398	-0.000889	0.216400
2001-04-02	1145.87	-0.012462	-0.002430	-0.000967	0.216743
2001-04-03	1106.46	-0.034393	-0.002566	-0.001123	0.221166

Figure 2. Dataset and Engineered Features

A more sophisticated RL Trading Agent

We now define a state-based model that will help us make our trading decision of going long or short. We had designed a set of features (Returns, Moving Averages, Rolling Volatility…) for our stock: for each feature, we define a Z-score, and the state is defined as the tuple of those Z-scores.

We estimate the transition probabilities, and the associated rewards, that we define as being the averages of the associated returns times our position (long, short, neutral). The actions that can we can take at each step is going long, short or neutral.

As our actions does not influence the next state, our model is off-policy, and we can estimate the best action to take for each state based on our data.

A robust testing approach

We use a K-fold Cross-Validation methodology that is well adapted to the time dependency of our data: each test set corresponds to a time window succeeding the training set to avoid lookahead bias.

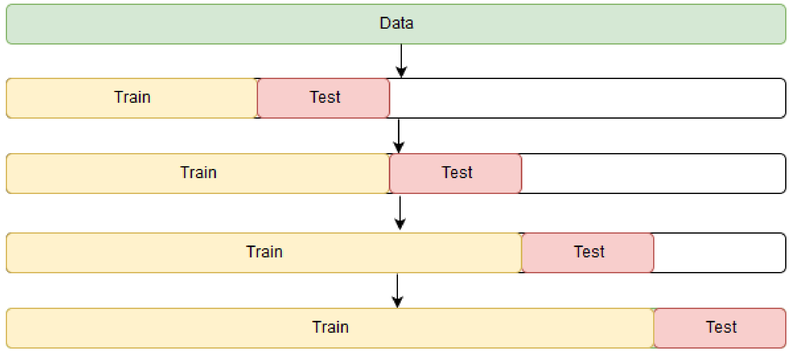

Figure 3. Cross Validation Methodology

For each test set (one year of data), we compute several financial ratio (Sharpe, Calmar, Max DD) to assess the performance of each model, and compare the average out of sample ratios to determine if one model is better than another.

	Sharpe			
	M1	M2	M3	M4
Test Set 1	0	0.32	0.38	0.41
Test Set 2	0	0.28	0.25	0.34
Test Set 3	0	0.54	0.6	0.98
Test Set 4	0	0.07	0.11	0.14
Test Set 5	0	0.13	0.21	0.25
Average	**0**	**0.268**	**0.31**	**0.42**

Figure 4. Out of Sample Sharpe Ratios Comparison

Conclusion

We manage to show that the Moving Averages are good signals to forecast stock trends, and that results are even better with a regularized prediction model. Our state-based model performs even better: we can think that using a Z-score acts as a regularization process, and unable us to have a better granularity on what is happening than trying to predict the return directly.