Reinforcement Learning for Trading cryptocurrencies

Rachel Duan

Katherine Wang

Yen Hann Yoo

INTRODUCTION

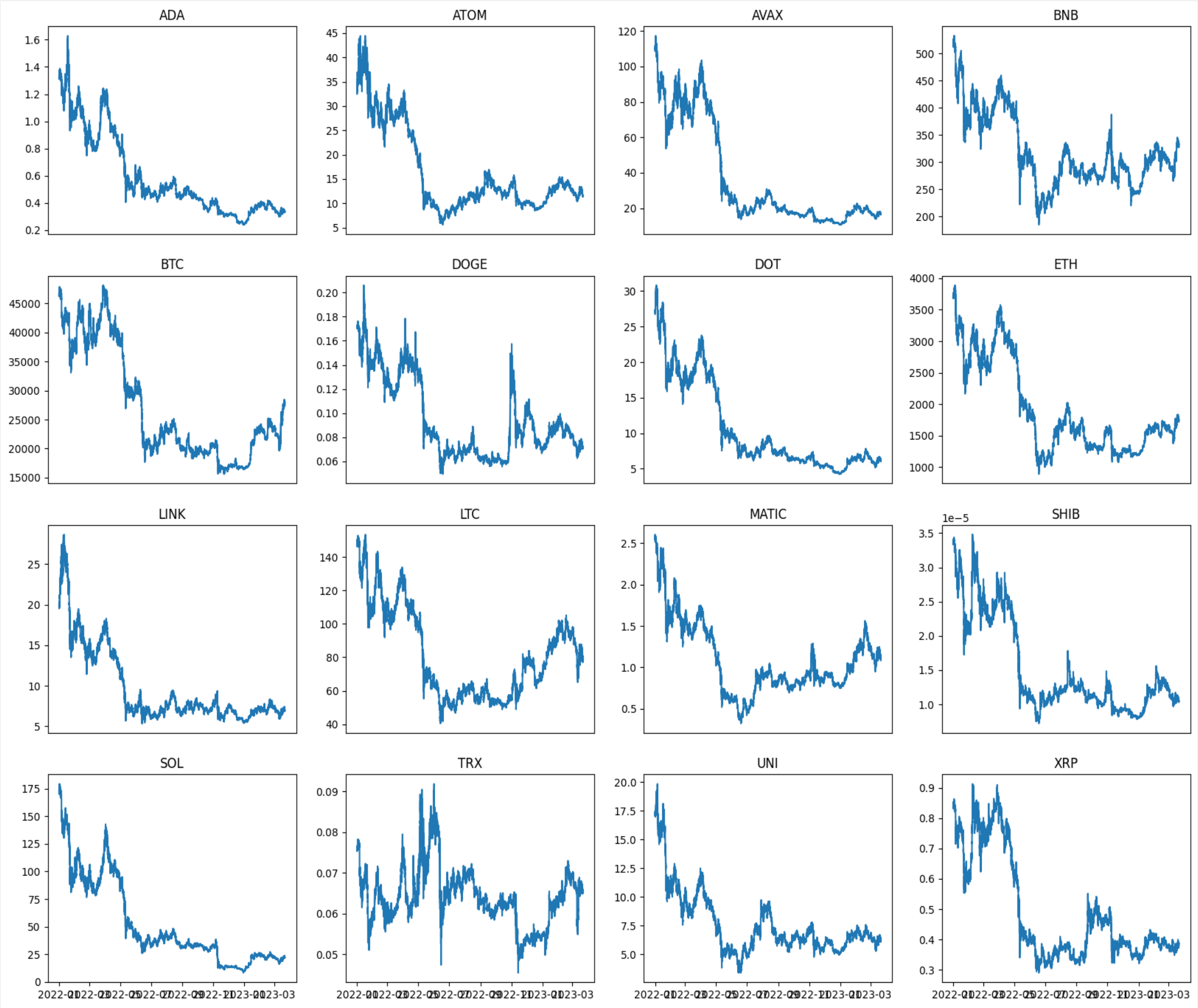
Motivation

- In financial markets, the noise to signal ratio is very low
- Even with the benefit of hindsight, it is often hard to tell what the best policy is
- RL allows us to learn the best trading policy without a model to predict price movement
- Automating trading reduces emotional biases and enables efficient portfolio management

Data

The data contains the following information for every 5 minutes from 01/01/2022 to 03/19/2023 for each of the 16 cryptocurrencies:

- Open price
- Close price
- Open time
- Close time
- High price
- Low price
- Volume
- Token



METHODOLOGY | STATE

Old Environment

Features
Close Prices
Periodic Returns over 1-hour

Current Environment

Features
Close Prices
Exponential Moving Average (EMA12)
Moving Avg. Conv. Div. (MACD)
Relative Strength Index (RSI14)
Rolling Volatility
Periodic Returns
Trading Volume

METHODOLOGY | ACTION & OBJECTIVE

Actions: A vector of floats over 16 cryptocurrencies

Weights: Action transformation - percentage of our portfolio value assigned to each token

Reward: Per period profit / loss based on weights

Old Environment

Only allow long (i.e., the action space has only positive values, between 0 and 1)

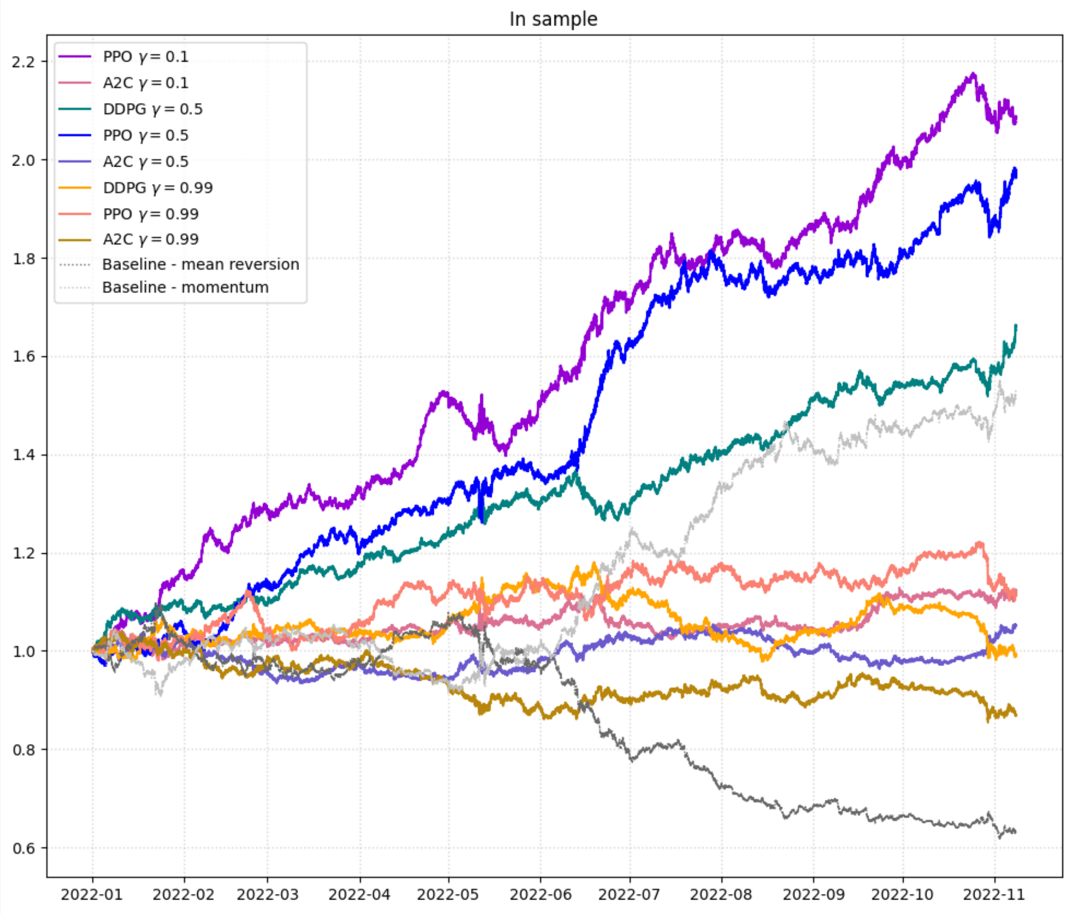
Weights are calculated using a softmax function that sums up to 1

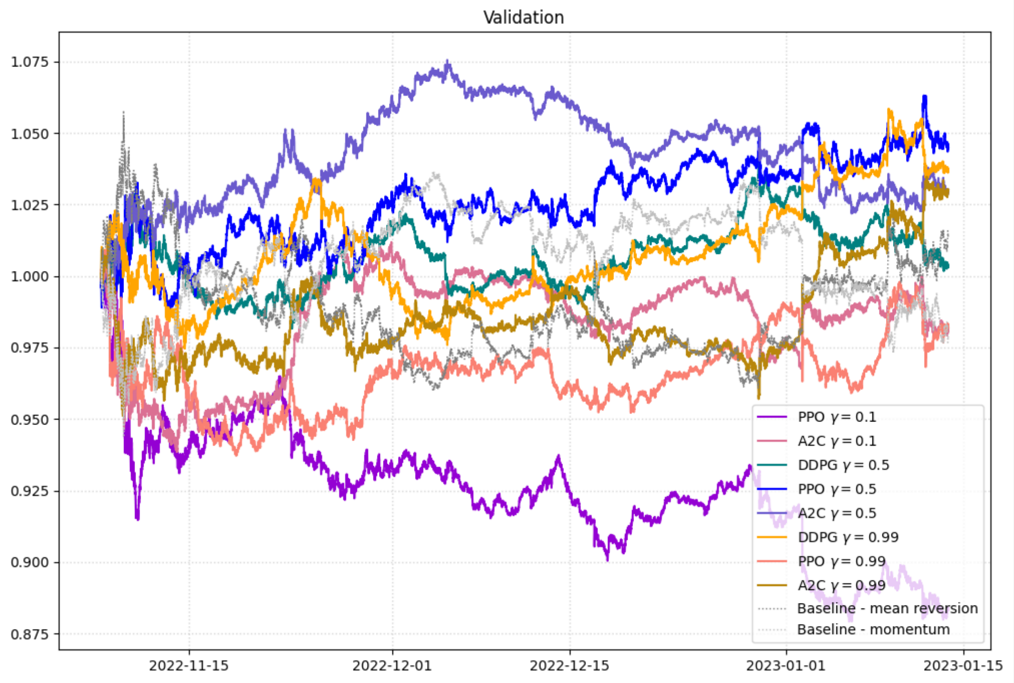
Current Environment

Allow long and short (i.e., the action space has both positive and negative values, between -1 and 1)

Weights are designed to sum up to 0, and absolute values of weights sum up to 1

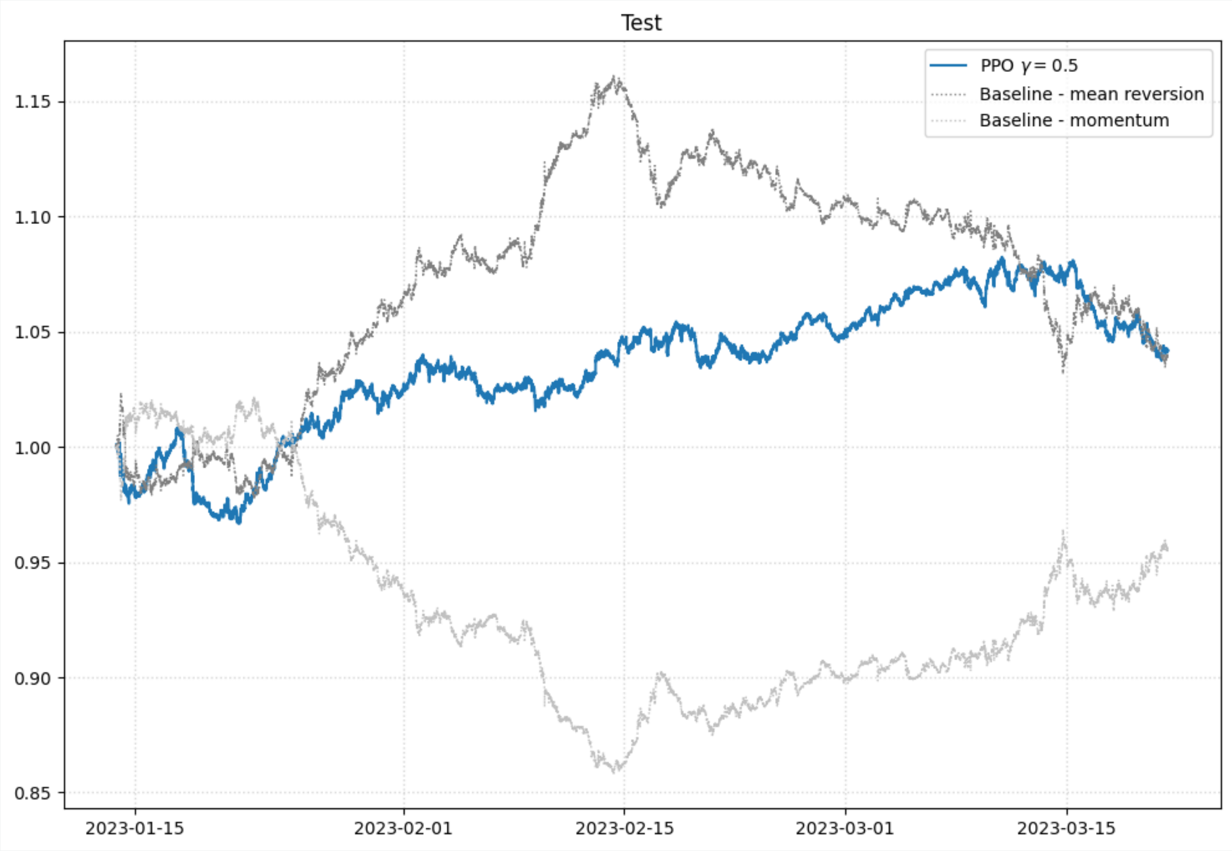
TRAIN / VALIDATION RESULTS





Validation Sharpe Ratio	A2C	PPO	DDPG
γ=0.1	-0.64	-3.44	
γ=0.5	1.80	**2.30**	0.47
γ=0.99	1.08	-0.51	1.68

TEST RESULTS



Our strategy Test Sharpe

1.63

Baseline Mean Reversion Sharpe

1.46

Baseline Momentum Sharpe

-1.46

Thank You

(Now you too can get rich quick)