

Order Placement Optimization with Reinforcement Learning

Research Question:

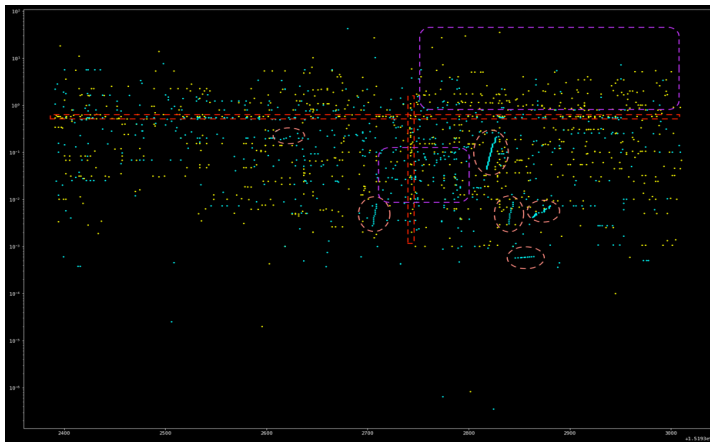
“How should one design a reinforcement learning environment and construct features, which are derived from a limit order book, in order to optimize limit order placement?”.

Order Placement Problem:

“How can we buy (respectively sell) V shares within a time horizon T , in a manner that minimizes the capital spent (respectively, maximizes the revenue received)?”.

Feature Engineering

Market events from traders



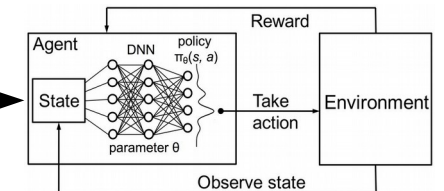
Order book

5724.00	5.378	18.488
5733.20	0.023	13.110
5731.50	0.003	13.087
5731.00	1.000	13.084
5730.10	0.088	12.084
5730.00	8.117	11.996
5729.00	0.873	3.879
5728.30	0.002	3.006
5728.00	2.308	3.004
5726.80	0.255	0.696
5726.20	0.008	0.441
5725.90	0.009	0.433
5725.00	0.424	0.424
5721.20 EUR		
5721.00	10.790	10.790
5720.10	14.420	25.210
5720.00	7.447	32.657
5718.90	3.210	35.867
5716.80	3.090	38.957
5715.60	1.448	40.405
5715.20	6.890	47.295
5714.30	0.050	47.345
5711.10	0.438	47.783
5711.00	0.013	47.796
5706.00	2.175	49.971
5697.40	0.260	50.231
5686.70	0.050	50.281

Feature: window of order book

$$S_{bidask} = \begin{bmatrix} bp_{11} & bs_{11} & bp_{21} & bs_{21} & \dots & bp_{m1} & bs_{m1} \\ bp_{12} & bs_{12} & bp_{22} & bs_{22} & \dots & bp_{m2} & bs_{m2} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ bp_{1n} & bs_{1n} & bp_{2n} & bs_{2n} & \dots & bp_{mn} & bs_{mn} \\ ap_{11} & as_{11} & ap_{21} & as_{21} & \dots & ap_{m1} & as_{m1} \\ ap_{12} & as_{12} & ap_{22} & as_{22} & \dots & ap_{m2} & as_{m2} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ ap_{1n} & as_{1n} & ap_{2n} & as_{2n} & \dots & ap_{mn} & as_{mn} \end{bmatrix}$$

Input for Deep Q-Network



Reinforcement Learning Setup

