

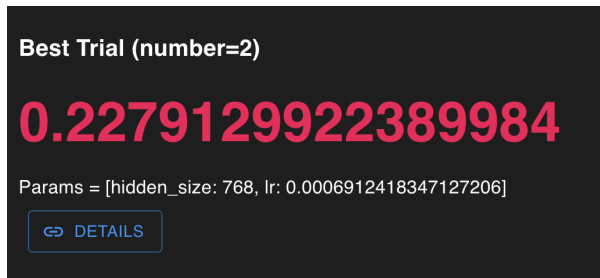
Assignment2PEF_AY23_24

Group 14

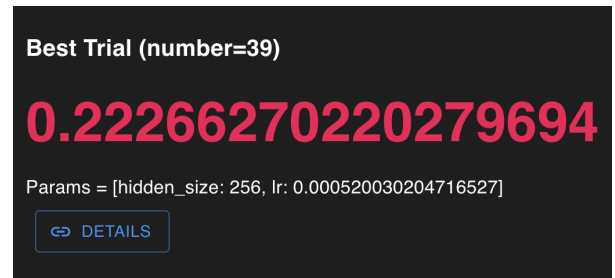
Idda, Raffaeli, Riondato, Stillo

1 Random Hyperparameter

Our purpose was to find the optimal hyperparameters of the gradient descent rule, i.e. the ones that minimize the loss function. To do so, we have chosen a random test setup and run the search on DNN with 10 and 50 trials. Increasing the number of trials could bring better results in terms of optimized hyperparameters, as the algorithm will have a better chance of finding combinations of hyperparameters that improve model performance. The discovered values are as follows:



(a) n_trials=10



(b) n_trials=50

EM_priceDNN-point-random_50

Number	State	Value	Param hidden_size	Param lr
0	COMPLETE	0.23271165788173676	512	0.001490985842573223
1	COMPLETE	0.26065072417259216	512	2.855790679080029e-05
2	COMPLETE	0.22891955077648163	576	0.0008054707437189927
3	COMPLETE	0.7948184013366699	128	0.09122314691019641
4	COMPLETE	0.2447117567062378	704	0.00010036986667204709
5	COMPLETE	0.2479686737060547	576	5.9075453080038385e-05
6	COMPLETE	0.23946700990200043	320	0.01437874626893263
7	COMPLETE	0.2448846399784088	704	3.1585619223070555e-05
8	COMPLETE	0.8141405582427979	768	0.04016235941640559
9	COMPLETE	0.7920532822608948	320	0.06023215781676299
10	COMPLETE	0.24019427597522736	832	0.0038918884613221554
11	COMPLETE	0.2608911395072937	704	1.897641769358223e-05
12	COMPLETE	0.24809348583221436	704	0.01172657974692062
13	COMPLETE	0.23795364797115326	192	0.0004980023593179826
14	COMPLETE	0.7711986303329468	896	0.044719765848669726
15	COMPLETE	0.2594795525074005	128	0.02363376805517803
16	COMPLETE	0.7065157294273376	576	0.029819972203391134
17	COMPLETE	0.24626068770885468	768	4.555008072794493e-05
18	COMPLETE	0.22481055557727814	256	0.0036769419779233677
19	COMPLETE	0.2521635890007019	960	1.4474775331160664e-05
20	COMPLETE	0.7599701881408691	576	0.029384702810435866
21	COMPLETE	0.2307414412498474	704	0.0007603043026803127
22	COMPLETE	0.8169125318527222	832	0.04163276326556236
23	COMPLETE	0.2399032711982727	384	0.005627612406115842
24	COMPLETE	0.2427065670490265	704	0.00010037493107644612
25	COMPLETE	0.2667083442211151	192	0.0001437505844321607
26	COMPLETE	0.23118770122528076	896	0.003734714529077203
27	COMPLETE	0.2374291867017746	512	0.0008412837752345288
28	COMPLETE	0.2526159882545471	512	4.4793171668043383e-05
29	COMPLETE	0.30403050780296326	832	0.021189860184112085
30	COMPLETE	0.23250731825828552	704	0.0015851587699397644
31	COMPLETE	0.7817985415458679	512	0.05457202930827474
32	COMPLETE	0.2651110589504242	256	7.101083534109171e-05
33	COMPLETE	0.2742108106613159	128	1.2472480501168678e-05
34	COMPLETE	0.23097358644008636	320	0.0002671320395447725
35	COMPLETE	0.23088231682777405	64	0.001409198448678033
36	COMPLETE	0.8038359880447388	576	0.05736439522213361
37	COMPLETE	0.22821485996246338	192	0.00044722570463238346
38	COMPLETE	0.22632160782814026	320	0.0005067343891391742
39	COMPLETE	0.22266270220279694	256	0.000520030204716527
40	COMPLETE	0.8092074394226074	320	0.05836803008421114
41	COMPLETE	0.2532581090927124	704	2.2625484146015587e-05
42	COMPLETE	0.8152645230293274	448	0.03779516531440767
43	COMPLETE	0.23800113797187805	384	0.006265441353605417
44	COMPLETE	0.24990324676036835	256	0.013217631831398178
45	COMPLETE	0.2763495147228241	64	0.00011466171759686724
46	COMPLETE	0.22679725289344788	832	0.0004930930267007044
47	COMPLETE	0.23833614587783813	128	0.001154111853219306
48	COMPLETE	0.22804538905620575	960	0.007786606919482678
49	COMPLETE	0.22617673873901367	704	0.001360054094991232

(b) n_trials=50

EM_priceDNN-point-random_10

Number	State	Value	Param hidden_size	Param lr
0	COMPLETE	0.2752843499183655	128	1.9944694760531313e-05
1	COMPLETE	0.23348571360111237	768	0.002719402203520313
2	COMPLETE	0.2279129922389984	768	0.0006912418347127206
3	COMPLETE	0.7164462804794312	576	0.022443033981790237
4	COMPLETE	0.28741654753685	128	0.04272041787833153
5	COMPLETE	0.2404654622077942	576	0.0001413731332191407
6	COMPLETE	0.2514691948890686	64	0.024564692989973696
7	COMPLETE	0.23557034134864807	192	0.003857085883050148
8	COMPLETE	0.25652015209198	896	1.4862853490173438e-05
9	COMPLETE	0.2820867598056793	64	4.3632591477522707e-05

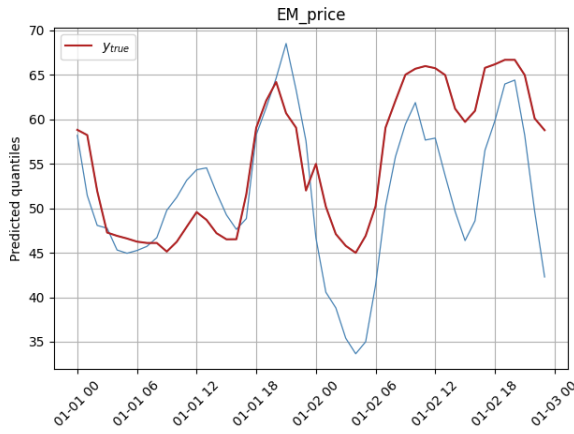
(a) n_trials=10

We can see that with more experiments the loss value becomes smaller, but is it enough? Indeed, we cannot rely solely on hyperparameters, as there are other factors to consider. For example, as observed in the table below, computational time increases consistently (by a factor of 5) while we increment `n_trials`.

	Computation time
<code>n_trials=10</code>	54.9 seconds
<code>n_trials=50</code>	254.8 seconds

Table 1: Computation time

Another feature to keep in mind is the approximation of the `EM_price`. We can observe that the approximation of the actual value of the price, from 01-02 00 until the end, is worse with 50 trials.



(a) `n_trials=10`



(b) `n_trials=50`

N.B. When we changed the number of `n_trials`, we paid attention to delete the previous one from the dashboard, to prevent the accumulation of experiments. Moreover, using a random test set every time we run the code the results were slightly different.