

Support Material: Grid Search for Feed Forward Neural Network

The following tables show the mean accuracy obtained after a gridsearch (with cross validation, $k = 5$) using Keras Classifier. If not specified otherwise, the fixed parameters are epochs=150, batch size=50, $\eta = 0.1$, $\lambda = 2.4 \cdot 10^{-4}$ and the relu activation function in the one hidden layer.

Table 1 yields the best accuracy result of 0.686 using 'optimizer': 'Adagrad'. From thereon Adagrad is used as the solver. Table 2 yields the best accuracy result of 0.687 using 'epochs': 50, 'learn_rate': 0.1. Table 3 yields the best accuracy result of 0.677 using 'init_mode': 'uniform'. From thereon, uniform initialization of the kernel is used. Table 4 yields the best accuracy result of 0.666 using 'activation': 'sigmoid'. Table 5 yields the best accuracy result of 0.677 using 'dropout_rate': 0.1, 'weight_constraint': 2. Table 6 yields the best accuracy result of 0.685 using 'batch_size': 60, 'epochs': 50, 'neurons': 20.

Table 1: Results (mean accuracy and standard deviation) of gridsearch using Keras Classifier varying the type of gradient descent solver

optimizer	Mean	STD. Dev
SGD	0.678	0.0245
RMSprop	0.668	0.0266
Adagrad	0.686	0.0292
Adam	0.677	0.0164

Table 2: Results (mean accuracy and standard deviation) of gridsearch using Keras Classifier varying the learning rate and number of epochs

epochs	learn_rate	Mean	STD. Dev
10	0.001	0.494	0.0477
10	0.01	0.6	0.0294
10	0.1	0.663	0.0146
50	0.001	0.539	0.0425
50	0.01	0.618	0.0389
50	0.1	0.687	0.018
100	0.001	0.563	0.0347
100	0.01	0.623	0.0339
100	0.1	0.676	0.0195
150	0.001	0.521	0.0353
150	0.01	0.665	0.0155
150	0.1	0.683	0.027
200	0.001	0.584	0.00959
200	0.01	0.67	0.0318
200	0.1	0.667	0.0206

Table 3: Results (mean accuracy and standard deviation) of gridsearch using Keras Classifier varying the kernal initilization

init_mode	Mean	STD. Dev
uniform	0.677	0.027
lecun_uniform	0.665	0.018
normal	0.661	0.0217
zero	0.6	0.0303
glorot_normal	0.656	0.00911
glorot_uniform	0.672	0.0289
he_normal	0.664	0.0147
he_uniform	0.663	0.0276

Table 4: Results (mean accuracy and standard deviation) of gridsearch using Keras Classifier varying the activation function

activation	Mean	STD. Dev
softmax	0.638	0.0161
relu	0.658	0.0237
tanh	0.612	0.0333
sigmoid	0.666	0.0219
linear	0.566	0.0354

Table 5: Results (mean accuracy and standard deviation) of gridsearch using Keras Classifier adding dropout rate and weight constraint

dropout_rate	weight_constraint	Mean	STD. Dev
0	1	0.671	0.0325
0	2	0.661	0.0136
0	3	0.667	0.0349
0	4	0.667	0.0207
0	5	0.665	0.0231
0.1	1	0.667	0.0236
0.1	2	0.677	0.0273
0.1	3	0.661	0.0333
0.1	4	0.657	0.0253
0.1	5	0.676	0.0343
0.2	1	0.667	0.0263
0.2	2	0.674	0.0221
0.2	3	0.665	0.0199
0.2	4	0.67	0.0324
0.2	5	0.665	0.0202
0.3	1	0.668	0.0221
0.3	2	0.673	0.0206
0.3	3	0.669	0.026
0.3	4	0.673	0.0265
0.3	5	0.672	0.013
0.4	1	0.674	0.0214
0.4	2	0.672	0.0277

continued on next page

Table 5: Lange Tabelle mit Logtable Fortsetzung

dropout_rate	weight_constraint	Mean	STD. Dev
0.4	3	0.663	0.0263
0.4	4	0.672	0.0163
0.4	5	0.667	0.0235
0.5	1	0.675	0.0219
0.5	2	0.675	0.0174
0.5	3	0.671	0.0308
0.5	4	0.672	0.0264
0.5	5	0.672	0.0104
0.6	1	0.661	0.0306
0.6	2	0.658	0.0288
0.6	3	0.663	0.0295
0.6	4	0.668	0.0241
0.6	5	0.672	0.0225
0.7	1	0.653	0.0376
0.7	2	0.653	0.0291
0.7	3	0.653	0.0174
0.7	4	0.666	0.0257
0.7	5	0.651	0.0264
0.8	1	0.637	0.0215
0.8	2	0.642	0.0374
0.8	3	0.648	0.0345
0.8	4	0.651	0.0153
0.8	5	0.643	0.0318
0.9	1	0.624	0.0295
0.9	2	0.632	0.0305
0.9	3	0.633	0.0342
0.9	4	0.623	0.0363
0.9	5	0.625	0.0372

Table 6: Results (mean accuracy and standard deviation) of gridsearch using Keras Classifier varying batch size, epochs and neurons

batch_size	epochs	neurons	Mean	STD. Dev
10	10	1	0.616	0.0298
10	10	5	0.652	0.0398
10	10	10	0.67	0.0221
10	10	15	0.673	0.0234
10	10	20	0.667	0.0211
10	10	25	0.668	0.0197
10	10	30	0.678	0.0251
10	10	100	0.671	0.0266
10	50	1	0.598	0.0282
10	50	5	0.656	0.0253
10	50	10	0.66	0.0238
10	50	15	0.67	0.0304
10	50	20	0.667	0.0169
10	50	25	0.668	0.0177

continued on next page

Table 6: Gridsearch varying batch size, epochs and neurons

batch_size	epochs	neurons	Mean	STD. Dev
10	50	30	0.657	0.0163
10	50	100	0.658	0.0239
10	100	1	0.612	0.0268
10	100	5	0.648	0.0514
10	100	10	0.671	0.0283
10	100	15	0.667	0.0337
10	100	20	0.659	0.0346
10	100	25	0.677	0.0366
10	100	30	0.666	0.0349
10	100	100	0.654	0.029
10	150	1	0.6	0.0303
10	150	5	0.663	0.0352
10	150	10	0.658	0.0184
10	150	15	0.672	0.0206
10	150	20	0.675	0.0299
10	150	25	0.674	0.022
10	150	30	0.665	0.0239
10	150	100	0.656	0.0346
20	10	1	0.6	0.0303
20	10	5	0.63	0.0288
20	10	10	0.65	0.0344
20	10	15	0.672	0.0131
20	10	20	0.673	0.0283
20	10	25	0.675	0.0263
20	10	30	0.672	0.0209
20	10	100	0.665	0.0276
20	50	1	0.623	0.0356
20	50	5	0.643	0.0349
20	50	10	0.66	0.02
20	50	15	0.677	0.0208
20	50	20	0.67	0.0296
20	50	25	0.661	0.0324
20	50	30	0.669	0.0258
20	50	100	0.658	0.0187
20	100	1	0.614	0.0279
20	100	5	0.658	0.0357
20	100	10	0.66	0.0254
20	100	15	0.661	0.029
20	100	20	0.667	0.0241
20	100	25	0.674	0.0346
20	100	30	0.658	0.0214
20	100	100	0.658	0.0315
20	150	1	0.636	0.0248
20	150	5	0.679	0.0218
20	150	10	0.658	0.0286
20	150	15	0.667	0.027
20	150	20	0.667	0.0226
20	150	25	0.662	0.0344
20	150	30	0.658	0.0396

continued on next page

Table 6: Gridsearch varying batch size, epochs and neurons

batch_size	epochs	neurons	Mean	STD. Dev
20	150	100	0.655	0.0207
40	10	1	0.601	0.0295
40	10	5	0.624	0.0324
40	10	10	0.628	0.0221
40	10	15	0.645	0.0288
40	10	20	0.653	0.0216
40	10	25	0.656	0.0344
40	10	30	0.662	0.0197
40	10	100	0.676	0.0324
40	50	1	0.609	0.0412
40	50	5	0.664	0.00896
40	50	10	0.671	0.0314
40	50	15	0.674	0.0223
40	50	20	0.669	0.024
40	50	25	0.663	0.0251
40	50	30	0.672	0.0359
40	50	100	0.661	0.0346
40	100	1	0.624	0.0313
40	100	5	0.64	0.0102
40	100	10	0.667	0.0242
40	100	15	0.67	0.02
40	100	20	0.657	0.0372
40	100	25	0.672	0.0286
40	100	30	0.664	0.0232
40	100	100	0.672	0.0194
40	150	1	0.616	0.0329
40	150	5	0.647	0.0283
40	150	10	0.667	0.0155
40	150	15	0.668	0.023
40	150	20	0.666	0.029
40	150	25	0.677	0.0227
40	150	30	0.66	0.0238
40	150	100	0.68	0.0227
60	10	1	0.604	0.0329
60	10	5	0.602	0.0321
60	10	10	0.614	0.0416
60	10	15	0.609	0.0391
60	10	20	0.63	0.0129
60	10	25	0.623	0.0314
60	10	30	0.624	0.0314
60	10	100	0.66	0.00789
60	50	1	0.604	0.0341
60	50	5	0.644	0.0366
60	50	10	0.672	0.0147
60	50	15	0.663	0.021
60	50	20	0.685	0.0324
60	50	25	0.67	0.0297
60	50	30	0.682	0.0199
60	50	100	0.676	0.0214

continued on next page

Table 6: Gridsearch varying batch size, epochs and neurons

batch_size	epochs	neurons	Mean	STD. Dev
60	100	1	0.611	0.0222
60	100	5	0.672	0.0274
60	100	10	0.674	0.0298
60	100	15	0.663	0.0223
60	100	20	0.657	0.0223
60	100	25	0.658	0.0237
60	100	30	0.67	0.0279
60	100	100	0.666	0.0202
60	150	1	0.611	0.0258
60	150	5	0.661	0.0166
60	150	10	0.669	0.0195
60	150	15	0.658	0.0101
60	150	20	0.667	0.0263
60	150	25	0.669	0.0267
60	150	30	0.663	0.0243
60	150	100	0.66	0.0303
80	10	1	0.6	0.0303
80	10	5	0.6	0.0303
80	10	10	0.604	0.0334
80	10	15	0.604	0.0312
80	10	20	0.604	0.0341
80	10	25	0.613	0.0409
80	10	30	0.607	0.0329
80	10	100	0.632	0.0261
80	50	1	0.613	0.0379
80	50	5	0.645	0.0446
80	50	10	0.672	0.0194
80	50	15	0.663	0.0192
80	50	20	0.679	0.0195
80	50	25	0.667	0.0177
80	50	30	0.663	0.02
80	50	100	0.665	0.0413
80	100	1	0.607	0.0387
80	100	5	0.632	0.0469
80	100	10	0.666	0.0295
80	100	15	0.662	0.0269
80	100	20	0.673	0.0278
80	100	25	0.663	0.0219
80	100	30	0.659	0.0352
80	100	100	0.661	0.0301
80	150	1	0.626	0.0198
80	150	5	0.667	0.0286
80	150	10	0.676	0.03
80	150	15	0.682	0.0277
80	150	20	0.67	0.0204
80	150	25	0.658	0.0351
80	150	30	0.666	0.0384
80	150	100	0.667	0.0146