

Terrain Discrimination

Hanna Olbert, Lukas Kyzlik

18. 12. 2019

Section 1

Perceptron Classifier

Perceptron – benchmark

Perceptron(max_iter=100000)

Average training time [s]	69.76
Average score	0.542

Table: Average results from 10 iterations of the test.

Conclusion

0.542 is benchmark we need to beat with more advanced methods to justify their use.

Section 2

Multy-Layer Perceptron (MLP) Classifier

MLP L-BFGS

Limited-memory BFGS is optimization algorithm in the family of quasi-Newton methods. It should perform the best for small datasets like ours.

In the following tests we use this default values for MLP classifier (if some of them are changed, it is noted):

Solver	L-BFGS
Alpha	10^{-5}
Hidden layer sizes	1 layer, 15 neurons
Random state	fixed seed
Activation function	ReLU

MLP L-BFGS – Activation functions

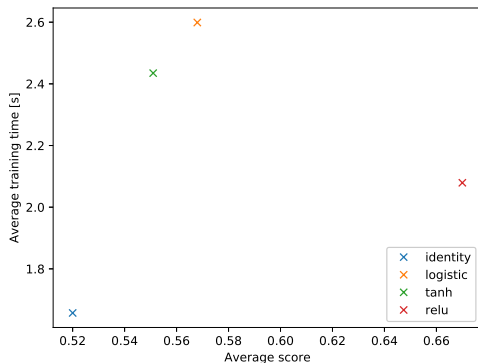


Figure: Average results from 20 iterations with different activation functions.

MLP L-BFGS – Activation functions

Conclusion

ReLU activation function gives the best results and is faster than most other functions for this problem.

MLP L-BFGS – Number of neurons (1 hidden layer)

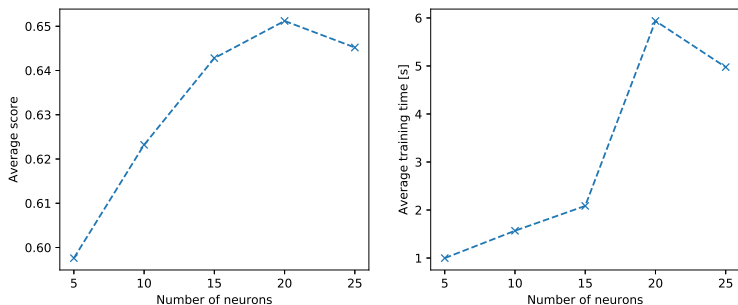


Figure: Average results from 50 iterations with different numbers of neurons in one hidden layer.

MLP L-BFGS – Number of neurons (1 hidden layer)

Conclusion

Optimal number of neurons with one hidden layer from perspective accuracy is around 20 for this problem. However, it also has the longest training time. Higher numbers of neurons probably result in over-fitting.

MLP L-BFGS – Number of neurons (2 hidden layers)

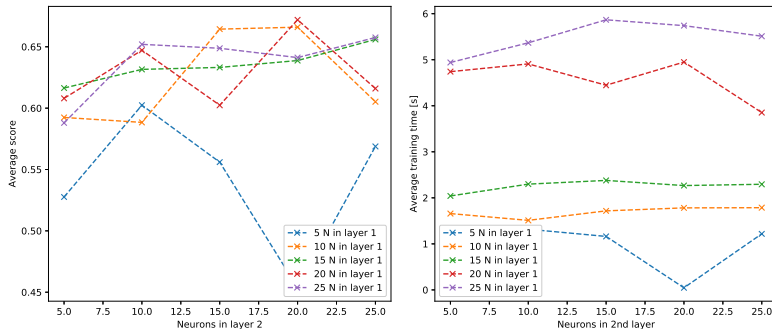


Figure: Average results from 50 iterations with different numbers of neurons in one hidden layer.

MLP SGD

Stochastic gradient descent is optimization algorithm for optimizing differentiable or subdifferentiable functions.

In the following tests we use this default values for MLP classifier (if some of them are changed, it is noted):

Solver	SGD
Alpha	10^{-5}
Hidden layer sizes	1 layer, 15 neurons
Random state	fixed seed
Activation function	ReLU
Batch size	200
Initial learning rate	0.001

MLP SGD – Initial learning rate

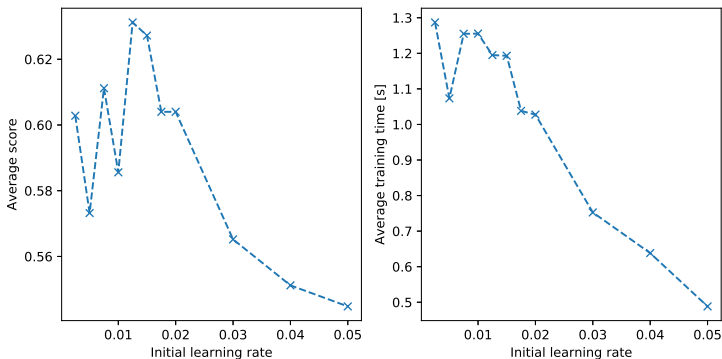


Figure: Average results from 50 iterations with different initial learning rates.

MLP SGD – Batch size

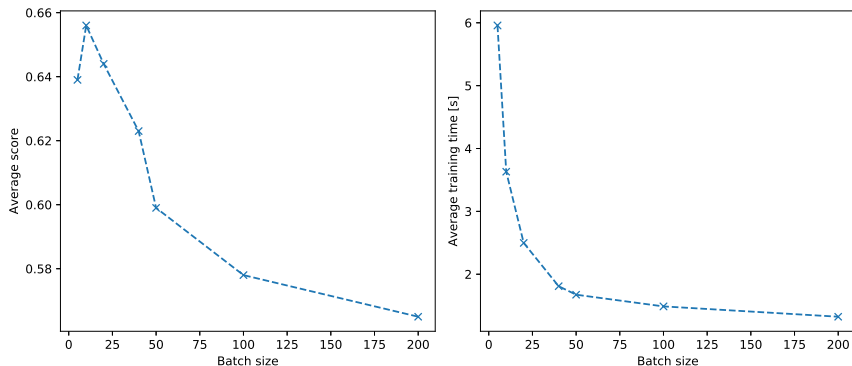


Figure: Average results from 20 iterations with different batch sizes.

