

Tuning parameters for prefix

Marinò Giosuè Cataldo¹

¹Dipartimento di Informatica, Università degli Studi di Milano,
Via Celoria 18 Milano, 20133, Italy

December 2, 2019

1 Model

In this tests we used a feed-forward Neural Network with 64 input neurons and 1 output neuron. The fixed settings are:

- The activation function of output neuros is Leaky-ReLU ($\alpha = 0.05$)
- Optimizer = SGD with Momentum ($\mu = 0.9$)
- Initializer for weights connection and bias is Random Normal with $\mu = 0$ and $\sigma = 0.05$
- Stopping Criteria: the mean absolute error don't improve with patience 10 or number of epochs is 20000
- Fixed Learning Rate

We try to find a better loss function than common Mean Square Error (MSE). So we try also the Mean Absolute Error (MAE) with many learning rate (η) and batch size (mb) on the three dataset we have.

2 Results

Table 1: File3

MSE			MAE		
$\eta = 1.0 \times 10^{-4}$					
mb	ϵ	Loss	mb	ϵ	Loss
16	8	3.16×10^{-5}	16	9	4.45×10^{-3}
32	8	3.15×10^{-5}	32	9	4.48×10^{-3}
64	8	3.15×10^{-5}	64	9	4.37×10^{-3}
$\eta = 1.0 \times 10^{-3}$					
16	8	3.15×10^{-5}	16	14	7.63×10^{-3}
32	8	3.15×10^{-5}	32	11	4.94×10^{-3}
64	8	3.15×10^{-5}	64	11	6.5×10^{-3}
$\eta = 5.0 \times 10^{-3}$					
16	8	3.16×10^{-5}	16	60	2.34×10^{-2}
32	8	3.16×10^{-5}	32	33	1.71×10^{-2}
64	8	3.15×10^{-5}	64	20	9.64×10^{-3}
$\eta = 1.0 \times 10^{-2}$					
16	8	3.22×10^{-5}	16	57	2.96×10^{-2}
32	8	3.34×10^{-5}	32	83	3.91×10^{-2}
64	8	3.15×10^{-5}	64	65	6.11×10^{-2}
$\eta = 5.0 \times 10^{-2}$					
16	11	6.54×10^{-5}	16	434	2.05×10^{-1}
32	8	3.22×10^{-5}	32	181	9.2×10^{-2}
64	8	3.25×10^{-5}	64	194	9.93×10^{-2}
$\eta = 1.0 \times 10^{-1}$					
16	44	9.45×10^{-4}	16	533	2.88×10^{-1}
32	8	3.24×10^{-5}	32	917	4.37×10^{-1}
64	8	3.49×10^{-5}	64	383	2.68×10^{-1}

Table 2: File7

MSE			MAE		
$\eta = 1.0 \times 10^{-4}$					
mb	ϵ	Loss	mb	ϵ	Loss
16	43	3.48×10^{-6}	16	74	2.47×10^{-3}
32	42	3.48×10^{-6}	32	49	1.55×10^{-3}
64	43	3.48×10^{-6}	64	47	1.56×10^{-3}
$\eta = 1.0 \times 10^{-3}$					
16	42	3.49×10^{-6}	16	135	4.86×10^{-3}
32	43	3.49×10^{-6}	32	131	4.03×10^{-3}
64	43	3.48×10^{-6}	64	102	3.59×10^{-3}
$\eta = 5.0 \times 10^{-3}$					
16	45	3.51×10^{-6}	16	740	2.09×10^{-2}
32	42	3.48×10^{-6}	32	640	2.31×10^{-2}
64	42	3.49×10^{-6}	64	605	3.35×10^{-2}
$\eta = 1.0 \times 10^{-2}$					
16	46	3.57×10^{-6}	16	1080	3.29×10^{-2}
32	44	3.52×10^{-6}	32	1493	3.69×10^{-2}
64	42	3.49×10^{-6}	64	839	3.94×10^{-2}
$\eta = 5.0 \times 10^{-2}$					
16	54	4.15×10^{-6}	16	7958	2.83×10^{-1}
32	47	3.67×10^{-6}	32	4365	1.33×10^{-1}
64	56	3.85×10^{-6}	64	3848	1.51×10^{-1}
$\eta = 1.0 \times 10^{-1}$					
16	203	4.73×10^{-5}	16	11421	3.74×10^{-1}
32	75	1.01×10^{-5}	32	7154	2.28×10^{-1}
64	56	5.84×10^{-6}	64	8192	6.01×10^{-1}

Table 3: File10

MSE			MAE		
$\eta = 1.0 \times 10^{-4}$					
mb	ϵ	Loss	mb	ϵ	Loss
16	613	3.36×10^{-8}	16	3120	6.22×10^{-4}
32	616	3.36×10^{-8}	32	1805	3.1×10^{-4}
64	620	3.36×10^{-8}	64	1242	2.24×10^{-4}
$\eta = 1.0 \times 10^{-3}$					
16	610	3.45×10^{-8}	16	1.55×10^4	2.88×10^{-3}
32	616	3.36×10^{-8}	32	1.640×10^4	3.39×10^{-3}
64	626	3.36×10^{-8}	64	1.175×10^4	2.39×10^{-3}
$\eta = 5.0 \times 10^{-3}$					
16	629	3.52×10^{-8}	16	9.410×10^4	1.75×10^{-2}
32	617	3.46×10^{-8}	32	1.048×10^5	3.31×10^{-2}
64	610	3.52×10^{-8}	64	5.29×10^4	1.33×10^{-2}
$\eta = 1.0 \times 10^{-2}$					
16	691	3.49×10^{-8}	16	2.129×10^5	3.88×10^{-2}
32	622	3.51×10^{-8}	32	1.620×10^5	2.9×10^{-2}
64	635	3.46×10^{-8}	64	1.245×10^5	2.41×10^{-2}
$\eta = 5.0 \times 10^{-2}$					
16	847	5.23×10^{-8}	16	1.153×10^6	1.91×10^{-1}
32	788	3.85×10^{-8}	32	6.299×10^5	1.1×10^{-1}
64	647	4.22×10^{-8}	64	3.650×10^5	7.41×10^{-2}
$\eta = 1.0 \times 10^{-1}$					
16	2884	4.09×10^{-7}	16	1.048×10^6	6.79×10^{-1}
32	909	4.45×10^{-8}	32	1.183×10^6	2.29×10^{-1}
64	984	7.45×10^{-8}	64	1.049×10^6	5.66×10^{-1}