

Dynamics of the Learning Process of a Multilayer Neural Network when using the AdaDelta Optimization Method

Monday 11 September 2023 14:20 (20 minutes)

Abstract. —The AdaDelta learning process optimization method has been tested for a multilayer neural network with three hidden layers with 28 neurons each, when recognizing printed numbers. Testing the learning error of this neural network was carried out using the mapping function and Fourier spectra of the error function. The mapping function describes the process of doubling the number of local minima. It is found that the application of the AdaDelta optimization method leads to a radically different picture of the behavior of the branching diagram than when applying the learning optimization methods Adam, AdamMax, and AMSGrad. Namely, with an increase in the number of iterations, the process is “extinguished” by retraining, by correcting the learning step of each neuron. It is shown that the hyperparameter ρ , which describes the contribution of the gradient square of the error function, significantly influences on the learning process of the neural network. That is, increasing the optimization parameter ρ to $(0.9 \div 0.999)$ causes a decrease in the region of change of the magnitude of the learning step. The step is selected automatically during training, for each neuron according to the AdaDelta optimization algorithm. It is shown that the optimal value of the optimization parameter ρ is 0.9 .

Author: Prof. SVELEBA, Sergij (Ivan Franko National University of Lviv)

Co-authors: Dr KATERYNCHUK, Ivan (Ivan Franko National University of Lviv); Dr KUNO, Ivan (Ivan Franko National University of Lviv); Dr BRYGILEVYCH, Volodymyr (The State Higher School of Technology and Economics in Jarosław)

Presenter: Dr KATERYNCHUK, Ivan (Ivan Franko National University of Lviv)

Session Classification: Computational Intelligence in Engineering

Track Classification: Computational intelligence in engineering