Can Machine Learning Help in Approximating Continuous Functions Satisfactorily

Irina Perfilieva[0000−0003−1531−1111]

IRAFM, University of Ostrava, 30. dubna 22, 701 03 Ostrava, Czech Republic
irina.perfilieva@osu.cz
http://www.osu.eu/

Based on the stunning success of neural networks in various fields, we believe in their ability to solve virtually all problems in science and applications. On the other hand, many researchers call for finding out what can explain or justify this success.

We contribute to the theoretical framework of neural networks (NNs) by focusing on the issue of the relationship between NN architecture, parameter selection, feed-forward computing capabilities, and learning from a limited data set.

We analyze neural networks with one hidden layer, known in the literature as single-layer feedforward neural networks (SLFN). NNs of this type are usefully used to solve practical problems, such as image processing, speech recognition, and systems control.

It is well known that the SLFN computational models implement a set of functions characterized as a weighted sum of ridge functions. This fact imposes the problem of describing sets of real functions that can be approximated by weighted sums of ridge functions. The solution to this problem provides with the knowledge of the approximation abilities of SLFNs.

The most influential results are associated with the names of G. Cybenko [1], K. Hornik [2], G.-B. Huang [3] (This list is by no means complete). These authors analyzed possible settings of SLFN parameters to prove that the function classes they define are dense in certain classes of smooth real functions of many variables. For example, in 1989 G. Cybenko proved that for any continuous sigmoidal activation function, the weighted (finite) sums of ridge functions are dense in the set of continuous functions on the unit cube, [1]. Further results discussed other options for choosing SLFN parameters: continuous, bounded and non-constant activation functions [2], random assignment of input weights and biases [3]. At the same time, all authors emphasized that for a satisfactory approximation, a sufficient number of neurons in the hidden layer is necessary.

However, almost all authors mentioned above did not consider learning from examples in their papers. They referred to backpropagation (BP) as the most popular learning algorithm. In contrast to the backpropagation, a learning strategy called the Extreme Learning Machine (ELM) was proposed in [3].

This paper analyzes the capabilities and limitations of two learning strategies: BP and ELM. We have created an example dataset consisting of input-output pairs that are samples of some continuous function on a closed interval. Therefore, according to [1], there is a neural network that satisfactorily approximates

this function. However, we have shown that this dataset cannot be correctly computed using SLFN trained with ELM following the algorithm in [3]. Moreover, our various experiments on training SLFN on this dataset using BP were also unsuccessful. On the other hand, we can constructively prove that at least one SLFN correctly computes the outputs of our dataset. The latter confirms the theory from [1] but does not confirm the unlimited possibilities of the most popular learning methods.

Keywords: single-layer feedforward neural networks · Extreme learning machine · machine learning

References

1. Cybenko, G.: Approximation by superpositions of a sigmoidal function. Math. Control Signals Systems, 2/4, 303-314 (1989)
2. Hornik, K., Stinchcombe, M., White, H.: Multilayer feedforward networks are universal approximators. Neural Networks, 2, 359âĂŞ366 (1989)
3. Huang, G.B.,Zhu, Q.Y., Siew, C.K.: Extreme learning machine: theory and applications. Neurocomputing, 70, 489âĂŞ501 (2006)