World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:17, No:04, 2023

An Empirical Study on Switching Activation Functions in Shallow and Deep Neural Networks

Authors : Apoorva Vinod, Archana Mathur, Snehanshu Saha

Abstract : Though there exists a plethora of Activation Functions (AFs) used in single and multiple hidden layer Neural Networks (NN), their behavior always raised curiosity, whether used in combination or singly. The popular AFs –Sigmoid, ReLU, and Tanh–have performed prominently well for shallow and deep architectures. Most of the time, AFs are used singly in multi-layered NN, and, to the best of our knowledge, their performance is never studied and analyzed deeply when used in combination. In this manuscript, we experiment with multi-layered NN architecture (both on shallow and deep architectures; Convolutional NN and VGG16) and investigate how well the network responds to using two different AFs (Sigmoid-Tanh, Tanh-ReLU, ReLU-Sigmoid) used alternately against a traditional, single (Sigmoid-Sigmoid, Tanh-Tanh, ReLUReLU) combination. Our results show that using two different AFs, the network achieves better accuracy, substantially lower loss, and faster convergence on 4 computer vision (CV) and 15 Non-CV (NCV) datasets. When using different AFs, not only was the accuracy greater by 6-7%, but we also accomplished convergence twice as fast. We present a case study to investigate the probability of networks suffering vanishing and exploding gradients when using two different AFs. Additionally, we theoretically showed that a composition of two or more AFs satisfies Universal Approximation Theorem (UAT).

Keywords : activation function, universal approximation function, neural networks, convergence
Conference Title : ICMLA 2023 : International Conference on Machine Learning and Applications
Conference Location : Boston, United States
Conference Dates : April 17-18, 2023