American International University- Bangladesh (AIUB)

Department of Computer Science

COMPUTER VISION AND PATTERN RECOGNITION [A]

Spring 2021-2022

Assignment Title: Review of Activation Functions Using Deep Learning Algorithm

Name	**ID**
Mahfuza Sharmili Jui	19-39398-1
Zafrin Sultana	19-39345-1
Maizul Islam	19-41002-2
Shoumik Saha	19-39348-1

SUBMITTED TO:

DEBAJYOTI KARMAKER

DATE OF SUBMISSION: **12.03.2022**

Review of Activation Functions Using Deep Learning Algorithm

Mahfuza Sharmili Jui
Zafrin Sultana
Maizul Islam
Shoumik Saha

Abstract:

In recent years, neural networks have grown in popularity as a solution to a variety of challenges. To cope with various sorts of challenges, many types of neural networks have been introduced.The basic purpose of every neural network, on the other hand, is to use a hierarchy of layers to transform non-linearly separable input data into more linearly separable abstract characteristics. Linear and nonlinear functions are combined in these layers.Improvements to the ReLu division amendment By upgrading the conventional convolution neural network, incorporating the nearby reaction standardization layer, employing the most extreme stacking, and so on, active work is offered. The actuation work is utilized to construct the altered convolution neural network structure model, using the CIFAR-10 informative collection as the neural network contribution for model preparation and assessment, using the Google profundity learning stage TensorFlow.We are analyzing the effects of the different Neuron enactment activities on the coexisting speed and image recognition precision.

Keywords- *Activation functions, ReLu, Non-linearity, tensorFlow, CNN, Deep neural network.*

Introduction:

Deep Learning (DL) is a contemporary topic of machine learning research that simulates the human cerebrum's functioning tool for breaking down and becoming familiar with images, music, language, and other information. Deep learning is a machine learning approach that is based on the representation of data. A convolution neural network and numerous other great machine learning techniques have grown as a result of the study and development of the deep learning strategy, and have won progress in different applications,for example, picture acknowledgment, target classification, etc [1].

A study of trainable activation functions in the neural network area was published , emphasizing general and particular properties of recent and previous techniques [4]. Trainable activation functions were studied in the context of feedforward neural networks in particular, while many of the techniques presented may also be used with recurrent neural networks. First and foremost, these methodologies' significant and crucial qualities are separated. A taxonomy is proposed based on this study that describes and classifies these functions according to their specification. Additionally, many of the proposed methods are demonstrated to be equivalent to adding neuron layers with fixed activation functions (non-trainable activation functions) and a simple local rule confining the weight layer.

The activation function is an important part of a convolution neural network that can plan nonlinear aspects in data. As a result, the convolution neural network has sufficient capacity to capture the complex case. The nonlinear components are added by the activation function, which removes unnecessary data while preserving highlights [1]. It stores "dynamic neuron highlights" and uses nonlinear processes to lead these highlights out. It is built on a neural network to address the difficult nonlinear challenge. The goal of this study is to look at how activation functions affect deep learning algorithms.

Literature Review:

In recent decades, ANNs have gained popularity as a result of successful applications in a variety of domains, including sediment movement, satellite image categorization, evapotranspiration, and coastal erosion. For complicated non-linear connections, they are a highly effective computing approach. There are at least three levels in their structure: input, hidden, and output. The coactive autonomic inference system model, recurrent network models, radial basis, multilayer perceptron, forward support vector regression, and other ANNs have all been created [3]. The backpropagation learning rule, which is an effective and practical learning method, has been applied to the multiple layer perceptron (MLP).Cortical neurons rarely occur at their maximal saturation regime, according the research, which suggests that their activation function may be represented by a rectifier.

That is to say, the neurons' functioning state is characterized by sparsity. Only one percent to four percent of the brain's neurons may be activated at the same time. However, over half of the neuron units in neural networks with sigmoid or logistic sigmoid activation functions are active at the same moment, which contradicts neuroscience studies. Furthermore, increasing the number of neuron units activated will make deep neural network training more complex [2]. Nair and Hinton first introduced Rectified linear units (ReLU) for Restricted Boltzmann machines.

Discussion:

Among the most popular activation functions have been discovered to be smooth and monotone. These characteristics are typically shared by hand-engineered activation functions [2]. Due to the fact that these activation functions do not trump ReLU, the fact that WRN-28-10 was able to attain such high accuracy with these arbitrary functions raises issues about what makes an activation function effective. The excellent accuracy gained by evolution in S2 indicates that sophisticated activation functions may rival with simpler form. Some instabilities are observed at NW using the logistic activation function. Although Parascandolo et al. [3], who recommended that the Tanh function might be substituted by the Sine function, our findings suggest that different functions could be appropriate to different scenarios or areas. It's possible that several counterintuitive activation functions exist that outperform more generic activation functions in particular contexts. Evolution is in a good position to find them.

The principal shows the performance of several activation functions Within the same domain however, it is possible to evolve with smaller systems and datasets and subsequently scale up to larger architectures and more complex datasets. By evaluating functions across many architectures and datasets, it may be possible to take such generalization even further in the future. The major feature of activation function meta-learning is that it enables you to find functions that are specific to each architecture and dataset. Most important gains are feasible in the environment.

Conclusion:

Deep neural networks have advanced rapidly in recent decades, particularly in computer vision and natural language processing. The saturated activation functions, like Sigmoid, hyperbolic tangent, are replaced by non-saturated counterparts, such as ReLU, ELU. Along with the network's layers, the power of search in huge spaces is demonstrated by evolution with loss-based fitness, which produces activation functions that reach high accuracy and outperform traditional functions such as ReLU and new functions such as Swish. Based on the conventional convolution neural network, the local response normalization layer, utilizes the maximum pooling, etc. Logistic and Tanh functions were shown to perform better than Identity, Exponential and Sine Functions. The primary strength of activation function meta-learning, on the other hand, is in discovering tailored functions for each architecture and dataset, resulting in significant improvements.

Contribution:

Paper Assessment

Student Name and ID	Paper No. from Ref	Paper Title
Islam, Maizul 19-41002-2	1	Review on Improved ReLu Piecewise Activation Function used in Deep learning Algorithms
Zafrin Sultana 19-39345-1	2	Activation Functions and Their Characteristics in Deep Neural Networks
Mahfuza Sharmili Jui 19-39398-1	3	Comparing Activation Functions in Modeling Shoreline Variation Using Multilayer Perceptron Neural Network
Shoumik Saha 19-39348-1	4	A survey on modern trainable activation functions

Paper Writing Contribution

Student Name and ID	Section No.	Section Title
Islam, Maizul 19-41002-2	1, 5	Abstract, Conclusion
Zafrin Sultana 19-39345-1	2	Introduction
Mahfuza Sharmili Jui 19-39398-1	3	Literature Review
Shoumik Saha 19-39348-1	4	Discussion

References:

1. Anusha Chitneni, Karunakar Pothuganti “Review on Improved ReLu Piecewise Activation Function used in Deep learning Algorithms” (November, 2018)
2. Bin Ding, Huimin Qian, Jun Zhou “Activation Functions and Their Characteristics in Deep Neural Networks” (2018)

3. Je-Chian Chen and Yu-Min Wang "Comparing Activation Functions in Modeling Shoreline Variation Using Multilayer Perceptron Neural Network" (April, 2020)

4. Andrea Apicella, Francesco Donnarumma, Francesco Isgrò, Roberto Prevete "A survey on modern trainable activation functions" (February, 2021)