QUIZ ASSESSMENT

Submitted By:
HARSHIT GAUR

MASTER OF PROFESSIONAL STUDIES IN ANALYTICS
ALY 6040 : DATA MINING
AUGUST 11, 2022

Submitted To: **PROF. KASUN SAMARASINGHE**

Question 1:

Importance of Activation Function in Neural Network

By generating a weighted total and then including bias with it, the activation function determines whether or not a neuron should be turned on. The activation function's objective is to add non-linearity to a neuron's output. It is utilized to determine the neural network's output, such as yes or no. The obtained values are mapped between 0 and 1 or -1 and 1, etc.

A neuron's activation status is determined by an activation function. By employing simpler mathematical procedures, it will determine whether or not the neuron's input to the network is significant during the prediction process. We are aware that neurons in neural networks behave in accordance with weight, bias, and their respective activation roles. The weights and biases of the neurons in a neural network would be updated based on the output error. Back-propagation is the name of this procedure. Back-propagation is made possible by activation functions since they provide the gradients and error needed to update the weights and biases.

The 3 types of activation functions in a neural network are:

1. ***Sigmoid or Logistic Activation Function***

   The Sigmoid Activation function, often known as the logistic function, has long been a favourite activation function for neural networks. The function converts the input into a number between 0.0 and 1.0. Values significantly lower than 0.0 are snapped to 0.0, and inputs substantially bigger than 1.0 are changed to the value 1.0. The function has an S-shape from zero up through 0.5 to 1.0 for all potential inputs. It was the standard activation used on neural networks for a considerable amount of time, up until the early 1990s.

2. ***Binary Step Function***

   Binary step function is a threshold-based activation function, meaning that once a specific threshold is reached, activation occurs, and below that point, deactivation occurs. The threshold is zero on the graph up top. As the name implies, this activation function can be utilized in binary classifications, but it cannot be employed when dealing with numerous classes.

3. ***Linear Activation Function***

   The Linear activation is proportionate to the input in a linear activation function, also referred to as "no activation" or the "identity function" (multiplied by 1.0). The function just spits out the value it was given, doing nothing to the weighted sum of the input.

Question 2:

Bagging Ensemble Method

Ensemble models (commonly referred to as "weak learners") are taught to tackle the same problem using the ensemble learning paradigm, which then combines the findings to produce better ones. The basic claim is that by properly combining weak models, we can produce more precise and/or reliable models.

This phenomenon occurs when all of the models are integrated to produce the best machine learning model. In such a way that we can use assistance from other models or combine assistance from all the models to produce the best hypothesis and outcomes when any one of the combined models begins to falter.

We fit the several learners separately from one another using parallel approaches, making it possible to train them simultaneously. The most well-known method of this type is "bagging," which stands for "bootstrap aggregating" and tries to create an ensemble model that is stronger than the individual models that make up it.

Whether we are working with a classification or regression problem, we receive a function during the training process that takes an input, returns an output, and is defined in relation to the training dataset. The fitted model is also subject to variability because of the theoretical variance of the training dataset (remember that a dataset is an observed sample coming from a genuine unknown underlying distribution); if another dataset had been observed, we would have received a different model.

The basic concept behind bagging is to fit a number of different models and "average" their forecasts to get a model with a lower variance. However, fitting totally independent models is not practical because it would necessitate

Procedure for Bagging are:

1. Consider a training set that contains m features and n observations. You must choose a random sample without replacement from the practice dataset.
2. Using sample data, a model is constructed using a subset of m features that is randomly selected.
3. The nodes are divided using the feature that offers the best split among all of them.
4. You have the best root nodes because the tree has matured.
5. Repeating the previous steps n times. To provide the most accurate prediction, it combines the results of various decision trees.