Short Answer Portion of Project 4

Emma Ozias and Peyton Skwarczysnki

1. Write about what an activation function is, and how they are used:

   The activation function is applied to the output of a neuron to introduce non-linearity to the model. When you chain together linear functions, all you get is another longer linear function. In order to solve more complex problems, activation functions are needed for the model to learn and behave differently than a linear regression model. These functions work by calculating the weighted sum and adding bias to it. This calculation then determines whether a neuron is activated.

2. Write about both of the following activation functions:

   a. Softmax: The softmax function is used for classification purposes. When doing classification, the sigmoid function can only handle two classes, but the softmax function can handle many classes. This activation function is important because it can tell you the probability that any of the many classes are true. This activation function works by taking in a vector of raw output generated by the previous nodes of the neural network and returns a vector of probability scores.

   b. Rectified Linear Unit Function (ReLU): ReLU is the most widely used activation function. It introduces nonlinearity to the model, and it helps reduce issues during gradient descent. All positive inputs are returned without modification, but the negative inputs are all returned as zero. This activation function is used as the default, since it has very fast computation time and generally works very well.

3. Pick any two of the following activation functions and research them, and then write a summary of what each does and how they are different.

a. Sigmoid (logistic): The sigmoid function is S-shaped, continuous, and differentiable. It is different from the tanh function because it has an output range of 0 to 1. The sigmoid function is used for binary classification, as seen through the limited output range. When looking at a binary classification model, the output layer often utilizes the sigmoid function. Despite its effectiveness at binary classification, this activation function suffers from the vanishing gradient problem.

b. Hyperbolic Tangent (tanh): This function is similar to the sigmoid function because it is also S-shaped, continuous, and differentiable. One important difference between them, however, is that the tanh function has an output range of -1 to 1. This expanded range causes each layer's output to be centered around 0 at the beginning of training, ultimately causing the function to be better at dealing with negative values. The tanh activation function is commonly used in the hidden layers of a deep neural network because of this zero-centered nature that enables fast computation speeds.

4. What is a Sequential model in the context of Keras?

   In Keras, there are two ways to build models: either sequential or functional. A sequential model refers to a plain, linear stack of layers. Each layer of this model only has a single tensor input and output. Sequential models are not used in scenarios that require multiple inputs or outputs, layer sharing, or a non-linear topology. Sequential models are beneficial due to their straightforward model structure, but are strictly limited in their complexity.

5. Write about each of the following layer types that can be used with neural networks:

a. Dense: Dense layers, also known as fully connected layers, are layers that are deeply connected with the previous layer. This dense layer will receive an input from each neuron in the previous layer, leading to each node in the dense layer having multiple inputs. In other words, all the neurons in one layer are connected to all the neurons in the following layer. Dense layers are the most commonly used layer in artificial neural networks.
b. Flatten: This type of layer is used to transform multi-dimensional input data into a one-dimensional array. Flatten layers are always used before dense layers. They take the output from the previous layer, transforms it into a one-dimensional array, then passes it to the dense layer. During the transformation process, all dimensions are collapsed except for the batch dimension. Flatten layers are important because they connect the convolutional and pooling layers with the dense layers of a convolutional neural network.

Works Cited

Ali, Moez. “Introduction to Activation Functions in Neural Networks.” *DataCamp,* Nov. 2023, www.datacamp.com/tutorial/introduction-to-activation-functions-in-neural-networks#.

Brownlee, Jason. “A Gentle Introduction to the Rectified Linear Unit (ReLU).” *Machine Learning Mastery*, 20 Aug. 2020, machinelearningmastery.com/rectified-linear-activation-function-for-deep-learning-neural-networks.

C, Bala Priya. “Softmax Activation Function: Everything You Need to Know.” *Pinecone,* 30 June 2023, www.pinecone.io/learn/softmax-activation/#The-Softmax-Activation-Function-Explained.

Chollet, François. “The Sequential Model.” *Keras,* 25 June 2023, keras.io/guides/sequential_model.

“Dense Neural Networks: Understanding Their Structure and Function.” *DataScientest,* 5 Mar. 2024, datascientest.com/en/dense-neural-networks-understanding-their-structure-and-function#:~:text=A%20dense%20neural%20network%20is,perform%20calculations%20on%20the%20data.

GeeksforGeeks. “Activation Functions in Neural Networks.” *GeeksforGeeks*, 16 July 2024, www.geeksforgeeks.org/activation-functions-neural-networks.

---. “Understanding Activation Functions in Depth.” *GeeksforGeeks*, 25 July 2024, www.geeksforgeeks.org/understanding-activation-functions-in-depth.

---. “What Is a Neural Network Flatten Layer?” *GeeksforGeeks,* 13 Feb. 2024, www.geeksforgeeks.org/what-is-a-neural-network-flatten-layer.

Géron, Aurélien. *Hands-On Machine Learning With Scikit-Learn, Keras, and TensorFlow*. 3rd ed., O’Reilly Media, 2022.