Assignment 06 Solutions

1. What is the difference between TRAINABLE and NON-TRAINABLE PARAMETERS ?

Ans. To sum-up: 'trainable parameters' are those which value is modified according to their gradient (the derivative of the error/loss/cost relative to the parameter), whereas 'non-trainable parameters' are those which value is not optimized according to their gradient.

2. In the CNN architecture, where does the DROPOUT LAYER go ?

Ans. Usually, dropout is placed on the fully connected layers only because they are the one with the greater number of parameters and thus they're likely to excessively co-adapting themselves causing overfitting.

3. What is the optimal number of hidden layers to stack ?

Ans. one rule says it should be 2/3 to the total number of inputs so if you 18 features , try 12 hidden neurons. For hidden layers two layers are considered better, however you can check the accuracy of your model by changing the numbers of layers and select the one which gives best result.

4. In each layer, how many secret units or filters should there be ?

Ans. it is common for a convolutional layer to learn from 32 to 512 filters in parallel for a given input.

5. What should your initial learning rate be ?

Ans. A traditional default value for the learning rate is 0.1 or 0.01, and this may represent a good starting point on your problem.

6. What do you do with the activation function ?

Ans. The activation function is a node that is put at the end of or in between Neural Networks. They help to decide if the neuron would fire or not.

7. What is NORMALIZATION OF DATA ?

Ans. Normalization is a technique often applied as part of data preparation for machine learning. The goal of normalization is to change the values of numeric columns in the dataset to use a common scale, without distorting differences in the ranges of values or losing information.

8. What is IMAGE AUGMENTATION and how does it work ?

Ans. Image augmentation is a technique of altering the existing data to create some more data for the model training process. In other words, it is

the process of artificially expanding the available dataset for training a deep learning model.

9. What is DECLINE IN LEARNING RATE ?

Ans. Perhaps the simplest learning rate schedule is to decrease the learning rate linearly from a large initial value to a small value. This allows large weight changes in the beginning of the learning process and small changes or fine-tuning towards the end of the learning process.

10. What does EARLY STOPPING CRITERIA mean ?

Ans. Early stopping is a method that allows you to specify an arbitrary large number of training epochs and stop training once the model performance stops improving on a hold out validation dataset. In this tutorial, you will discover the Keras API for adding early stopping to overfit deep learning neural network models.