

## Q1

It occurs when a model is too simple, which can be a result of a model needing more training time, more input features, or less regularization. Like overfitting, when a model is underfitted, it cannot establish the dominant trend within the data, resulting in training errors and poor performance of the model.

Overfitting occurs when the model cannot generalize and fits too closely to the training dataset instead. Overfitting happens due to several reasons, such as: The training data size is too small and does not contain enough data samples to accurately represent all possible input data values.

Fix overfitting:

1. Increasing the training data by data augmentation.
2. Feature selection by choosing the best features and remove the useless/unnecessary features.
3. Early stopping the training of deep learning models where the number of epochs is set high.

Underfitting problem can be solved by using a bigger network (more hidden nodes). A large number of nodes will be able to extract a good number of features thus understanding the data more efficiently. Also, you can train your model for a longer duration to avoid the underfitting problem.

## Q4

1-what's the effect of padding is equal same? What's another value for padding?

Using 'SAME' ensures that the filter is applied to all the elements of the input. Normally, padding is set to "SAME" while training the model. Output size is mathematically convenient for further computation and based on that, the output has the same size as the input

Another Value is none.

2- explain the affection of activation function.

Activation functions play an integral role in neural networks by introducing nonlinearity. This nonlinearity allows neural networks to develop complex representations and functions based on the inputs that would not be possible with a simple linear regression model.

3- explain the affection of using kernel\_initializer in layers.

The term “kernel initializer” is a fancy term for which statistical distribution or function to use for initializing the weights. In case of statistical distribution, the library will generate numbers from that

While building and training neural networks, it is **crucial** to initialize the weights appropriately to ensure a model with high accuracy. If the weights are not correctly initialized, it may give rise to the **Vanishing Gradient** problem or the Exploding Gradient problem.

#### 4- explain what's the difference between Conv2DTranspose and Conv2D.

Conv2D is mainly used when you want to detect features, e.g., in the encoder part of an autoencoder model, and it may shrink your input shape. Conversely, Conv2DTranspose is used for creating features, for example, in the decoder part of an autoencoder model for constructing an image

#### 5- explain downsample\_block, double\_conv\_block and upsample\_block functions.

Down sampling, which is also sometimes called decimation, reduces the sampling rate. Up sampling, or interpolation, increases the sampling rate. DCNNs is groups filters where filters within each group are translated versions of each other.

#### 6- why use an optimizer in learning?

We need an algorithm for minimize a loss function with respect to a model's trainable parameters. Based on that, we use optimizer function to change the attributes of your machine/deep learning model such as weights and learning rate in order to reduce the losses

#### 7- why use compile function?

We need to defines the loss function, the optimizer and the metrics. Based on that, we compile our model.

#### 8- why are we select categorical\_crossentropy in the loss of function?

We used as a loss function for multi-class classification model where there are two or more output labels; Based on that, we choose it.

#### 9 - explain early stopping function.

We used this function to stop our training when we reach a specific situation in our training process. In this code, this function monitor out "validation loss" value. After Enter into this situation, our model will be stopped and return its best wights based on "validation loss" value.

Also, we have another two parameters:

"Min delta" which is minimum change in the monitored quantity to qualify as an improvement, i.e., an absolute change of less than "Min delta", will count as no improvement.

"Patience": number of epochs with no improvement after which training will be stopped.

#### 10 - explain different between fit and compile functions in Keras.

By using compile command, our deep learning compilers take framework models as input and generate optimized codes for a variety of deep learning hardware as output; However, we use fit command for the model training on the data set for the specified number of fixed epochs or iterations mentioned.

So, compile will create our DL model and fit start out training process with that model.

#### 11- explain the difference between batch and epoch.

Epoch is the number of epochs is a hyperparameter that defines the number times that the learning algorithm will work through the entire training dataset and in the other hand, batch size is a hyperparameter that defines the number of samples to work through before updating the internal model parameters.

The batch size is a number of samples processed before the model is updated and the number of epochs is the number of complete passes through the training dataset.

The size of a batch must be more than or equal to one and less than or equal to the number of samples in the training dataset.

The number of epochs can be set to an integer value between one and infinity. You can run the algorithm for as long as you like and even stop it using other criteria besides a fixed number of epochs, such as a change (or lack of change) in model error over time.

They are both integer values and they are both hyperparameters for the learning algorithm, e.g. parameters for the learning process, not internal model parameters found by the learning process.

You must specify the batch size and number of epochs for a learning algorithm.

There are no magic rules for how to configure these parameters. You must try different values and see what works best for your problem.

Ref:

[Difference Between a Batch and an Epoch in a Neural Network - MachineLearningMastery.com](https://machinelearningmastery.com/difference-between-a-batch-and-an-epoch-in-a-neural-network/)