ML Assignment

Sarvesh S

21BDA05

**1.Have you come across Grid Search Cross Validation? Fit any two models covered in previous classes and optimize them using Grid search CV.**

Grid search is a process that searches exhaustively through a manually specified subset of the hyperparameter space of the targeted algorithm. Random search, on the other hand, selects a value for each hyperparameter independently using a probability distribution.

**Working procedure**:

Here we need to specify appropriate hyper parameter value of the parameter and the model name, and rest of the things needed for the parameter for eg: if we want to know grid search cv for random forest means we need to give probable numbers of trees we want, max depth of that and so on

It is used to give most appropriate method that can be used by our model to give the best

i.e statistically its like choosing a best estimator from a set of good estimators

**disadvantage:**

it takes more time for huge dataset(run time)

**syntax:**

clf = GridSearchCv(estimator, param\_grid, cv, scoring)

**2.What is Stride, Padding & Pooling? Explain with an example**

Explaining hyperparameter:

Stride:

Stride=1 which represent the filter window is moved by 1 pixel right and 1 pixel down, we can mention stride=2 or 3 and soon. Smaller strides lead to large overlaps which means the Output Volume is high. Larger strides lead to lesser overlaps which means lower output volume

Padding:

The first stage in cnn if we input image it will automatically convert that into form of array after we perform some convolution, stride and so on. In that while performing the corer value in that matrix is less participated in the feature extraction compared to inside to overcome this problem we use padding. Here we just adding one more row and column containing 0 so that inner values are covered, and one advantage is that after we do filter we can get the exact shape like the image we inputed

For eg: it is used in image detection if our object is present in the corner means through padding we can focus on corner too

Pooling:

Pooling is also known as down sampling that reduces the dimensionality but it retains the essential features. there are max, avg pooling

**4.What is overfitting? How to overcome overfitting in an ML model?**

overfitiing occurs when a statistical model fits exactly against its training data. When this happens, the algorithm unfortunately cannot perform accurately against unseen data.

Reduce the network's capacity by removing layers or reducing the number of elements in the hidden layers.

Apply regularization , which comes down to adding a cost to the loss function for large weights.

Use Dropout layers, which will randomly remove certain features by setting them to zero.