Test (April 21)

1. What is the purpose of having a ["training set" - "cross validation set" - "test set"] split of data in machine learning applications?

   Training Set is used to build up our algorithm. In this phase we usually create multiple algorithms in order to compare their performances during the Cross Validation Phase.

   Cross Validation Set gives us the insight on how the model will react to the unknown data. So we can find how good your model is trained. It is used as a parameter selection to avoid over-fitting.

   Test Set is used to test our prediction algorithm in order to see how it's going to perform so we can have an idea about our algorithm's performance on unseen data.

2. Suppose you did a linear regression fit on a training data using an nth degree polynomial and a set of p features. What procedure will you use to determine if over-fitting or under-fitting has occurred? (Assume you have training, cross-validation and test data sets in a 60-20-20 split).

   First of all we find J_{Train}, J_{CV} and J_{Test}.

   If J_{Train} is very low and J_{Test} is very high for more number of features or higher degree of polynomial, it is Over-Fitting.

   If J_{Train} is very high for less number of features or very lower degree of polynomial, it is Under-Fitting.

3. What procedure will you use to determine if collecting more training data is desirable to reduce test error in a regression model?

First select number of Polynomial and number of Features.

Fix the value of λ.

Find the appropriate θ values such as J_{Train} is low.

Calculate J_{Train}, J_{CV} and J_{Test}.

Check whether the model is high variance or high bias?

If the model is high variance, adding more data may help in reducing Test Error. But not for sure everytime.

And if the model is high bias, adding more data will never help to solve the problem.

4. What procedure will you use to determine if adding more features (alternatively adding polynomial features) will help reduce test error in a linear regression model?

First select number of Polynomial and number of Features.

Fix the value of λ.

Find the appropriate θ values such as J_{Train} is low.

Calculate J_{Train}, J_{CV} and J_{Test}.

Check whether the model is high variance or high bias?

If the model is high variance, adding more features will not help in reducing Test Error.

But if the model is high bias, adding more features will help to solve the problem.

5. Can you create a Neural Network that computes this function with (a) No hidden layers (b) One hidden layer? Show the network in the cases above where you can create one. Otherwise argue why such a network cannot be made.

For the given input and output, the Function is
$Y = (X_1 \text{ XOR } X_2) \text{ AND } (X_1 \text{ XOR } X_3)$

(a) No Hidden Layers

No, we cannot create a neural network for the above function without using hidden layers. For any combination of weights it is not possible to make this kind of network.

(b) One Hidden Layer

No, we cannot create a neural network for the above function with just One Hidden Layer because for the above function two processes take place

1. X_1 XOR X_2 and X_1 XOR X_3 at the same time
2. AND operation

XOR takes one hidden layer to complete and get the results, it cannot be done without hidden layer. Second hidden layer is the result of two XORs that does AND operation and computes the final result.

So, in any case for the above Function, it is not possible to create a Neural Network without a hidden layer or with just one hidden layer.

But it is possible to create it with two hidden layers.