REI602M **Quiz 12.3.2021**

Hand in on Gradescope before 22:00 on March 13 (Saturday). Each question will be given 1, 0.5 or 0 points as follows. If the question is more or less correct it gets 1 point. If it is partl y correct it gets 0.5, and if it is missing or completely wrong it gets 0 points.

1) [Exam April 2020] For each of the following questions, list all correct choices. Justify your answers briefly.

a) During the training of a convolutional neural network you observe that the training error does not decrease. What can you do to resolve the problem?

a) Change the network architecture;

Only as a last result since changing the network architecture can be very expensive.

b) Add dropout regularization;

c) Normalize the input data;

Reduce the data to normal forms to reduce data redundancy as well as improving data integrity.

d) Adjust the learning rate.

If adjusting the learning rate correctly, we can change how quickly the model converges to logacl minima. Best to lower the learning late once the model is close to local minima.

b) During the training of a neural network you observe that the training error has become very low but the validation error is fairly high. What can you do to resolve the problem?

a) Increase training data;

b) Add dropout regularization;

c) Perform data augmentation;

Helps with reducing overfitting

d) Add hidden layers while using the same training data.

2) Exam [June 2020] You are grading homework assignments in REI602M. In particular you are grading a problem where the students use a convolutional neural network to classify a large image data set. The students then have to compare their results with results obtained by others (available from a web page curated by the author of the data set). Assign Right / Wrong to each of the following statements given by the students and provide a one sentence explanation.

a) "My network outperforms all other classifiers on this data set as indicated by the extremely low training error."

Wrong:

Extremely low training error does not mean that this network has a good test error.

Let's say for example, the training data contains two unbalanced classes, yes and no, where yes occurs 99.5% of the time. Then the network would just classify all samples as yes and receive only 0.5% training error.

Let's say if the test data is classified as 99% No and 1 % Yes, the network would not be able accurately classify the training samples.

This is an example of **overfitting.**

b) "My network outperforms all other classifiers on this data set as indicated by the extremely low test error (results correspond to the best values of network hyperparameters)."

Right

hyperparameters determines the network structure, if the best parameters are chosen and there is a low test error. You have a good network.

c) "My network outperforms all other classifiers on this data set as indicated by the extremely low test error (I used dropout rate=0.1746562318)."

Wrong

The dropout value is very low, the dropout rate is between [0.0,1.0] where 1.0 means no dropout and 0.0 means no outputs from the layer. Where the dropout rate is low means we are not using enough data to be sure that the accuracy is correct, at least if you’re using the dropout in the testing phase. Drop rate is normally between

[0.5 – 0.8]