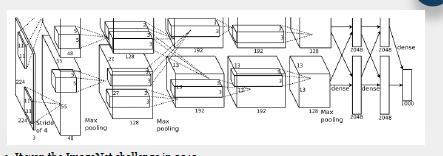
Advanced Machine Learning – week 5 - Regularization in CNNs and CNN based downstream tasks

Dr.Lei Zhang

Attendance code: 408773

[lzhang@Lincoln.ac.uk](mailto:lzhang@Lincoln.ac.uk)

CNN architecture-AlexNet in 2012-Recap



It won the ImageNet challenge in 2012.

AlexNet is considered to be the first CNN architecture which rose the interest in CNNs.

The net contains eight layers with weights; the first five are convolutional and the remaining three are full connected. The output of the last fully-connected layer is fed to a 1000-way SoftMax which produces a distribution over the 1000 class labels.

Softmax - The softmax function is **used as the activation function in the output layer of neural network models that predict a multinomial probability distribution**. That is, softmax is used as the activation function for multi-class classification problems where class membership is required on more than two class labels.

**CNN architecture – VGG16 2014 – Recap**

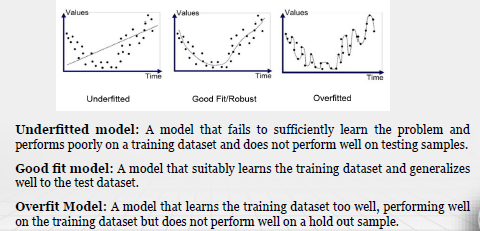
The main contribution is a thorough evaluation of networks of increasing depth using an architecture with very small (3 x 3) convolution filters, which shows that a significant improvement on the prior-art configuration can be achieved by pushing the depth to 16-19 weight layers.

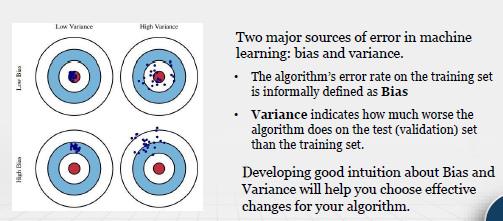
Learning Objectives:

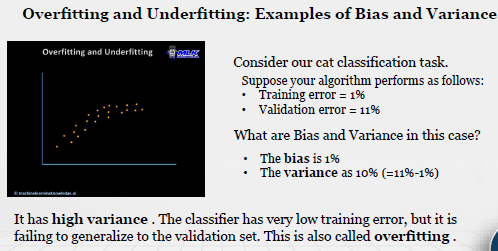
* To understand the overfitting and underfitting issues in ML
* To know the techniques to mitigate the learning problem and accelerate the training
* To know CNN based classic downstream tasks/applications

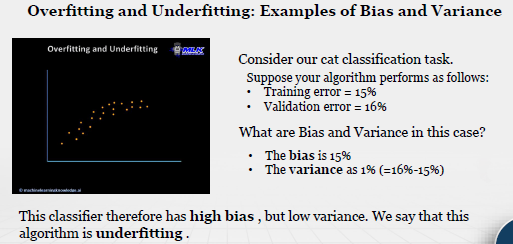
**Overfitting and Underfitting in Deep Learning Neural Networks**

*The central challenge in machine learning is that we must perform well on new, previously unseen inputs—not just those on which our model was trained. The ability to perform well on previously unobserved inputs is called generalization.*

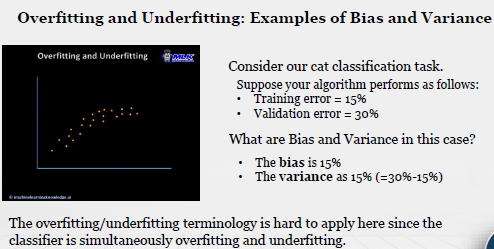


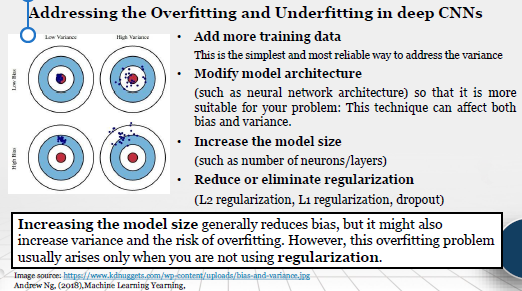






If bias is high this is underfitting, if the variance is high then this is overfitting.

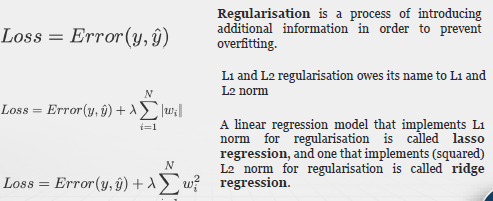


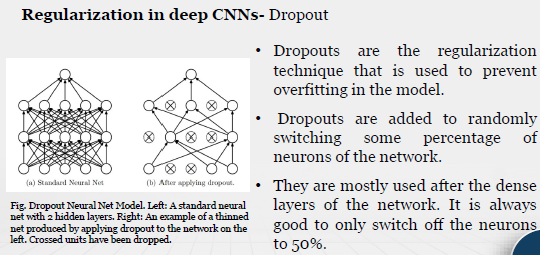


**Regularization in deep CNNs - L1 and L2 regularization**

The process of introducing additional information in order to prevent overfitting

L1 and l2 are used to regulate the loss function





Regularization in Deep CNNs – Batch Normalization

