Week 5 Practical

Dr Simon Denman
CAB420: Machine Learning

This weeks practical will focus on using fine tuning and data augmentation to adapt neural networks to new tasks using limited data.

Problem 1. Fine-Tuning DCNNs. The *Kuzushiji-MNIST-49* database, or K-MNIST-49, is a MNIST like database, though with a substantially larger number of classes. However while the number of classes and the variety within them is increased over MNIST, the characteristics of the images are largely the same, i.e. the database is made up of 28×28 pixel grey scale images, where pixel values are often close to 0 or 1 (i.e. black or white). Given this, a model trained on MNIST is likely to adapt well when fine-tuned on K-MNIST-49.

Using a model that has been trained on MNIST and the K-MNIST-49 dataset:

1. Modify the network to classify 49 classes rather than the original 10;
2. Fine-tune the network and assess the model's performance. Consider how it compares to the models trained from scratch in the previous week's practical.

Problem 2. Data Augmentation. The Houses dataset contains 535 sets of images of houses, and the corresponding price of those houses. There are a number of images for each house, with images covering the front, bathroom, kitchen and bedroom. Ordinarily, this would be too little data to train a deep neural network, however data augmentation offers one way to try to overcome this. Using this dataset, design and train a model to predict the house price from an image. In doing this you should:

1. Design a simple network for this task, bearing in mind that you have limited data. While you may wish to fine-tune from a dataset such as CIFAR, note that this will restrict you to images of size 32×32 [1].
2. Divide the dataset into appropriate training and testing splits.
3. Set appropriate data augmentation parameters to generate additional samples.
4. Train the network and evaluate it's performance. You may also which to consider which images to use. Using all images obviously leads to more data, but also increases the problem space, while using only (for example) frontal images may make the task easier as the network only needs to learn information relative to the front of the house.

[1] assuming you retain the input layer and all fully connected layers. If you change the input layer and remove all fully connected layers, you will be able to use a different input image size – note that this approach is not expected, but interested students are welcome to explore