Week 4 Practical

Dr Simon Denman
CAB420: Machine Learning

This weeks practical will focus on classification with deep neural networks.

Problem 1. Kuzushiji-MNIST and Kuzushiji-MNIST-49. The *Kuzushiji-MNIST* database is a MNIST like database based on Japanese Kuzushiji characters. Following the pattern of MNIST, it contains gresyscale images of size 28×28 pixels for 10 Kuzushiji characters. A second expanded dataset, *Kuzushiji-MNIST-49*, has the same format but 49 different characters. The dataset is also imbalanced, in that not every class has the same number of instances and some classes are rarer than others. Using these two datasets:

1. Create and train a network to classify the *Kuzushiji-MNIST* database into it's 10 classes. Evaluate the network on the test set.
2. Extend the network to classify the *Kuzushiji-MNIST-49* dataset into it's 49 classes. Evaluate the network on the test set, and consider the frequency with which each class occurs when analysing results.

Note that you do not need to try to make your networks overly complex, or train them for more than a few minutes. Smaller networks and training times of 5-10 epochs will be sufficient for this practical. However you are encouraged to experiment with different architectures and explore how they impact performance. In particular, you can explore the placement of:

- Batch Normalisation;
- Dropout;
- Different Activation Functions (i.e. Leaky Relu, ELU or Swish); as well as
- The parameters of convolutional and fully connected layers.