**University of Bristol Research Placement**

* 19th June – 25th August (10 weeks)
* ‘Use high resolution visual and thermal imagery of the turbine surfaces to identify cracks, delamination, and other damage using machine learning’
* Suggested techniques:
  + Super pixel saturation thresholding
  + Percolation models
  + Convolutional networks
  + Bayesian classifiers
  + Deep learning

My relevant experience:

* Machine vision work using Python
  + I built a pedestrian detection program about a month ago - this used a linear support vector machine to classify images according to their histograms of oriented gradients. Doing this involved using OpenCV, Theano, NumPy, and sklearn (Python libraries for image processing, numerical computing, and machine learning respectively). I did this in an attempt to replicate the results of a paper (‘Histograms of Oriented Gradients for Human Detection’, Dalal & Triggs).
  + I wrote an API for stitching Soviet-era maps together (for Val) – I used OpenCV again for this: it allows people to download hundreds of images of the maps, detect and crop their borders, correct perspective shift, stitch them together, fit gridlines to them, then export them at a resolution they think appropriate.
  + I’ve implemented convolutional neural networks (both my own simple ones and standard deeper ones such as VGG16) and multi-layer perceptrons to classify digits and common objects. I used Keras, Theano, and Tensorflow to do this, and trained the networks on GPU instances provided by AWS.
* Broader machine learning work in Python, Matlab, and R
  + In addition to what I describe above I’ve also put to use other supervised learning algorithms (e.g. linear and logistic regression, LDA, ensemble methods) through either readily available libraries or my own implementations. To build the latter I studied the prerequisite linear algebra through MIT OCW, and have been solving the problem sets made available through CS229 from Stanford.
  + Beyond supervised algorithms I know the common techniques for exploratory data analysis, dimensionality reduction, model shrinkage/subset selection, cross-validation, and unsupervised learning.
* Work at DCA
  + My ICR is built around applications I’ve written in Matlab and R. It will discuss (hopefully) an application-specific generalized additive model I’ve designed to help my team diagnose faults in a medical device based on the device’s displacement-load profile. I’m also looking at applications for bayesian statistics in DCA, and am drafting statistical production guidelines for one of our major medical clients.
  + DSP - I’ve created various convolutional filters to allow automatic processing measurements of a product’s sound and mechanical performance.
  + Statistical analyses of test results, usually comparing the expectation and variance of two groups.