

Automatic Curriculum Learning for Deep Models Using Active Learning -MSc Dissertation, Progress Report

Ian McWilliam s0904776

1 Background

The standard practice when training supervised deep learning models is to sample uniformly from a training set of examples, with each example having an equal probability of selection. There are two fields of research which offer alternatives to this paradigm; *curriculum learning* and *active learning*. Proponents of curriculum learning hypothesise that, analogously to how humans learn, it is beneficial to focus on learning from ‘easy’ examples, particularly early on in training, with various theoretical justifications such as comparisons to *continuation methods* and *transfer learning*. Contrastingly, many active learning approaches focus on finding examples about which the learning algorithm is most ‘uncertain’ about, and selecting those examples to continue training. The aim in both cases is to improve overall learning performance, by resulting in an algorithm with superior generalisation error, and/or one that converges to an optimal solution at a faster rate than uniform sampling or without needing as many training samples.

A key challenge in curriculum learning is determining how to measure the ‘difficulty’ of a given sample; in certain domains there is a clear delineation between ‘hard’ and ‘easy’ samples, however generally speaking this is not the case, and from our research there would appear to be limited literature on the subject as it pertains to curriculum learning. In active learning however there are a broad range of *acquisition functions* which can be used to select the next training example, and which we suggest can be used as a proxy for sample difficulty in a curriculum learning setting.

2 Project Goals

The goal of this project is to investigate the extent to which active learning methods can be used to enable automated curriculum learning in deep learning models, as well as investigating how the impact of the difficulty of training samples affect learning. The primary method through which we will construct an automatic curricula is the use a chosen active learning method (for example distance to classification threshold, expected model change, BALD method etc) to score each training sample. We then weight the sampling probability of each training sample proportionally to its score, leading to training examples with a higher sampling probability being selected significantly more often than others, biasing the learning process towards either ‘easy’ or ‘hard’ examples, depending on the chosen approach. We then train models on a particular dataset using this automatic curriculum method and compare their test performance to what would have been obtained if the models had been training using traditional uniform sampling; if the methods show consistent outperformance against the uniform sampling baseline we would conclude that our methods are successfully improving the learning process.

The current plan is to run experiments on three datasets:

- MNIST - Recognising hand-written digits.
- Geometric Shapes - Differentiating between rectangles, ellipses and triangles. With this dataset there is an easily defined curriculum where learning begins with regular shapes (i.e. squares, circles and equilateral triangles), providing another benchmark against which to compare the proposed methods’ performance.

- CIFAR 10/100 - Image recognition task.

We will test a variety of model architectures and examine how the automated curricula affect performance.

3 Progress and Next Steps

Currently, extensive experiments have been run on the MNIST and Geometric Shapes datasets, with encouraging results; in both datasets sampling with the the automatic curriculum methods leads to superior performance versus a uniform sampling benchmark on a held out test set, either in terms of higher classification accuracy or lower cross-entropy error. Analysis of the sampling probabilities allows identification of which samples are being selected more/less often in both datasets; in the Geometric Shapes dataset we observe that the distance to threshold metric successfully separates the easy, regular examples from the more difficult ones, resulting in an effect similar to the manually constructed curriculum. Similarly, in the MNIST examples, we observe that the curriculum successfully identifies shapes that are visually very difficult to classify, for example extremely idiosyncratic handwriting, with parts of the number appearing to be missing.

One issue encountered with the Geometric Shapes dataset was a disparity in performance between our results and those of the referenced paper, however this was discovered to be due to the chosen mini-batch size, as the paper appears to use stochastic gradient descent instead of mini-batch stochastic gradient descent. Rerunning the experiments with stochastic gradient descent resulted in similar relative performances, we therefore plan to use experiments using the mini-batch methods as we are able to run experiments significantly faster and therefore run a larger number of tests.

The next step in the project will be to begin running experiments on the CIFAR 10 and 100 datasets, as well as progress with writing up the report. As the testing infrastructure has already been developed for the first two datasets the testing of the CIFAR datasets should hopefully be relatively quick, particularly because we plan to use pre-trained networks which will accelerate training. Additionally, different active learning metrics will be explored as most testing has so far been done only with the distance to classification threshold metric; we will rerun experiments again with different ways of inferring sample difficulty and examine how the choice of difficulty metric affects learning.