Informatics Project Proposal - Automatic Curriculum Learning for Deep Models Using Active Learning

Ian McWilliam s0904776

Abstract

In this paper we propose a project researching how active learning methods can be used to automatically construct training curricula, and test the hypothesis that training deep models with such 'active curricula' can improve upon other training paradigms such as traditional curriculum learning, random sampling or pre-training. The work will investigate a variety of methodologies for constructing a training curriculum using active learning metrics, focusing on how the different training methods affect the performance of convolutional networks on a set of image classification tasks, with performance measured in terms of generalization error and convergence speed. The output of the project will hopefully be a set of novel ways to improve the training of deep models, as well as a more thorough exploration of the relationship between active and curriculum learning than currently exists in the literature.

1 Introduction

Active learning refers to the learning paradigm wherein machine learning algorithms actively select or 'query' the samples from which it learns; in the case of deep learning this contrasts with the default approach of randomly sampling 'batches' of labeled training samples. The motivation for active learning is that, by allowing it to select the samples from which it learns, the algorithm can achieve superior generalization performance from a smaller number of training samples than if the samples had been chosen randomly. In domains where unlabeled data is abundant but obtaining labels is expensive, active learning can be used to reduce the cost associated with training a deep model, as the active approach allows the designer to obtain labels only for the samples which will be most beneficial to learning.

There are a variety of methods by which the algorithm can query datapoints, however they generally focus on finding the points in the input space that the algorithm is most 'uncertain' about, allowing the algorithm to fill what could be seen as gaps in its knowledge of the domain. A related field is that of *curriculum learning*, which explores how the learning process can be improved by presenting training samples to the algorithm in a meaningful order (with the order defining a 'curriculum'). Motivated by the way in which humans and animals learn, the learning curriculum generally begins with 'easy' examples, with the difficulty of the examples increasing as training progresses; Bengio et al show that learning with a curriculum can improve overall performance on a variety of tasks etc. It is interesting to note that, while similar, active and curriculum learning are somewhat opposed in their learning philosophies, with the former focusing on learning from uncertain/difficult samples and the latter focusing beginning with easy samples before continuing to more difficult ones.

2 Purpose

DESCRIPTION OF PROPOSAL IDEA, I.E. CONSTRUCTING CURRICULA WITH ACTIVE LEARNING

3 Background

Outline key papers

4 Methods

Constructing curricula with a set of Active Learning methods. Various curriculum constructions, select easy, select hard, easy to hard, hard to easy

5 Evaluation

Compare to random sampling, pre-training. Use CNNs, maybe LSTMs too,

6 Outputs

New training paradigms, hopefully generic ways to consistently improve on random sampling

7 Workplan

References

[1] J.O'Keefe, J.Dostrovsky, "The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat", *Brain Research*, Issue 34, pp 171-175, 1971