BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI
HYDERABAD CAMPUS

FIRST SEMESTER 2015-2016
MATH F376 / MATH F377

Date: 24/1/2016

Application of Algorithmic Differentiation in Machine Learning

Need for Study
Machine Learning (especially Neural Networks, hereby referred to as ANN.) have become fundamental to the field of Artificial Intelligence. Deep Learning has risen as the prime candidate for high level, abstract learning representation. Training such ANNs using the famous backpropagation algorithm requires computation of several gradients. So far, these gradients have been calculated by hand, and then coded into the algorithm. This has caused ANNs to be limited in their expression. Algorithmic Differentiation would allow a larger variety of functions that can be represented by ANNs allowing for more diverse applications. It would also allow for reduced workload while prototyping new ANNs as gradients will not need to be manually calculated.

Objectives
Understand fundamental principles of Algorithmic Differentiation and ANN representation. Evaluate the requirements for creating a library tailored for the purpose of representing and training ANNs. (Factors include Computation and Space efficiency, Memory Management, Flexibility)
Create and test the library on some standard networks and datasets, such as, MNIST, CIFAR, IRIS etc.

Literature Review
Application of Algorithmic Differentiation in ANNs has been researched to great extent over the past year[2]. Some working libraries include Torch/autograd[4], Theano[5] and Tensorflow[6]. Torch has been shown to be 2x slower than a hand coded gradient implementation, while Tensorflow has been show to have lower performance rating than Torch[7].

Evaluating Derivatives[1] is a fundamental book on the principles of Algorithmic differentiation. The book follows a hybrid approach between forward and reverse mode. Which is suitable for our purpose of developing Machine Learning algorithms. The book also addresses programming related issues such as implementing evaluation trees with floating point values and vector mode calculation of partial derivatives. Given the complex nature of classification problems that need to be addressed by Deep Learning, the Library needs to support parallelisation via GPU enhancement.

Algorithmic Differentiation has just started being applied in the field of Machine Learning and its scope is tremendous. It allows for easier experimentation of ANN architecture and hence quickens prototyping process.

Work Plan

Date	Progress Checkpoint
25/2/16	Understanding of fundamental concepts of Algorithmic Differentiation, and associated Machine Learning flavors of AD.
8/3/16	Select appropriate language environment based on addressing of programming related issues. Abstract library design ready.
25/3/16	Standard ANN architecture functioning on benchmark public datasets. Submission of mid-sem report.
25/4/16	Implement successively complex features with compatibility for extension to diverse network architectures. Attempt GPU optimisation and parallelisation of computations. Refactor and document code. Make library usable. Submit final report.

References

[1] Evaluating Derivatives – Principles and Techniques of Algorithmic Differentiation (Andreas Griewank, Andrea Walther)
[2] Automatic Differentiation in Machine Learning: A Survey. (Baydin et al.)
[3] Automatic Differentiation of Algorithms in Machine Learning. (Baydin et al.)
[4] https://blog.twitter.com/2015/autograd-for-torch
[6] https://github.com/zer0n/deepframeworks/blob/master/README.md
[5] https://www.tensorflow.org/

Expected Knowledge Gain
- Strong concepts in Linear Algebra, Activation and Loss/Error Functions, Vector Calculus, Convexity Analysis.
- Ability to apply Algorithmic Differentiation to various gradient based task.
- Usage of Neural Networks for a large variety of problems such as Image Classification, Video Analysis, Language Models, Time Series Data etc.
- Understand concepts of Parallelisation and Scalability of Design.

Name: Chivukula Satya Adityakrishna
ID no: 2013A7PS387H