Machine Learning - CSCI 5622

Project Proposal

Exploring bayesian methods to improve neural network learning

Santhanakrishnan Ramani, Shane Grigsby[1], Shirly Montero, Shruthi Sukumar

Goal

The goal of this project is twofold :

(a) To optimize hyper-parameters of a neural network using Gaussian Processes based on the algorithm outlined in [1]. In addition to the best estimates of the hyper-parameters, the confidence in those parameters will be obtained via simulations from the GP.

(b) To improve neural network performance by considering the networks as probabilistic models and doing Bayesian inference.

Description

For the hyperparameter optimization of the neural network, we would focus on building an automatic tuning algorithm using Gaussian processes based on [1]. The Spearmint library is the first choice for the GP implementation. Different objectives for the GP will be explored and compared to the baseline tuning algorithms grid search and randomized search.

The field of Bayesian neural networks has been an active area of research since the early 1990's [2, 3]. A rapid advancement in Bayesian inference methods over the past decade, e.g. Slice Sampling in MCMC and variational inference, has renewed interest in the field. Goal (b) of the project would aim to explore the use of MCMC to sample weights for "shallow" neural networks, and perform variational inference to improve generalization of deep networks during training [4, 5]. These methods are to be compared to the backpropogation training algorithm as baseline.

Tools

(a) Spearmint library for Bayesian Optimization

(b) Keras for Neural network training

DataSet

(a) World Development Indicators Dataset link

(b) NASA ICESat/IceBridge dataset [6] [7] link

References

[1] Jasper Snoek, Hugo Larochelle, and Ryan P Adams. Practical bayesian optimization of machine learning algorithms. pages 2951–2959, 2012.

[2] Radford M Neal. Bayesian learning for neural networks. 118, 2012.

[3] David JC MacKay. Bayesian methods for adaptive models. 1992.

[4] Yarin Gal and Zoubin Ghahramani. On modern deep learning and variational inference.

[5] Alex Graves. Practical variational inference for neural networks. In *Advances in Neural Information Processing Systems*, pages 2348–2356, 2011.

[6] Zwally J. et al. (2002). Icesats laser measurements of polar ice, atmosphere, ocean and land, j. geodyn., 34, 405 445.

[7] Dotson Ryan C. and John Arvesen 2014. Icebridge dms l3 photogrammetric dem, [indicate subset used]. boulder, colorado usa: Nasa daac at the national snow and ice data center.

[1] Not enrolled in CSCI 5622