ASU School of Mathematical and Statistical Sciences
Arizona State University

MEMORANDUM

DATE: March 7, 2022

TO: Faculty and Students

FROM: Professor(s) Lalitha Sankar Shiwei Lan
Chair/Co-Chairs of Erika Cole
Defense for the MA in Mathematics
Committee Members Giulia Pedrielli
Paul Hahn

DEFENSE ANNOUNCEMENT

Candidate: Erika Cole

Defense Date: 04/13/2022

Defense Time: 2:30 PM

Virtual Meeting Link: https://asu.zoom.us/j/87628714917

Title: Bayesian Methods to Tune Hyperparameters of Loss Functions in Machine Learning

Please share this information with colleagues and other students, especially those studying in similar fields. Faculty and students are encouraged to attend. The defending candidate will give a 40 minute talk, after which the committee members will ask questions. There may be time for questions from those in attendance. But, guests are primarily invited to attend as observers and will be excused when the committee begins its deliberations or if the committee wishes to question the candidate privately.

ABSTRACT
-See next page-

PO Box 871804 Tempe, AZ 85287-1804
(480) 965-3951 Fax: (480) 965-8119
http://math.asu.edu

Abstract

The introduction of parameterized loss functions for robustness in machine learning has led to questions as to how hyperparameter(s) of the loss functions can be tuned. This thesis explores how Bayesian methods can be leveraged to tune such hyperparameters. Specifically, a modified Gibbs sampling scheme is used to generate a distribution of loss parameters of tunable loss functions. The modified Gibbs sampler is a two-block sampler that alternates between sampling the loss parameter and optimizing the other model parameters. The sampling step is performed using slice sampling, while the optimization step is performed using gradient descent. This thesis explores the application of the modified Gibbs sampler to alpha-loss, a tunable loss function with a single parameter $\alpha \in [0, \quad)$, that is designed for the classification setting. Theoretically, it is shown that the Markov chain generated by a modified Gibbs sampling scheme is ergodic; that is, the chain has, and converges to, a unique stationary (posterior) distribution. Further, the modified Gibbs sampler is implemented in two experiments: a synthetic dataset and a canonical image dataset. The results show that the modified Gibbs sampler performs well under label noise, generating a distribution indicating preference for larger values of alpha, matching the outcomes of previous experiments.