WPI MATHEMATICAL SCIENCES

Wenjing Li

PhD Candidate

Mathematical Sciences

Wednesday, September 30, 2020

2:00 PM-4:00 PM

For Zoom information, please contact:
rpodell@wpi.edu

Dissertation Committee:

Dr. Randy C. Paffenroth, WPI (advisor)
Dr. Xiangnan Kong, WPI
Dr. Oren Mangoubi, WPI
Dr. Jian Zou, WPI
Dr. Simon Tavener, Colorado State University

PhD Dissertation Proposal Presentation

Optimal Ensembles for Deep Learning Classification: Theory and Practice

Abstract:

Ensemble learning is a process by which multiple base learners are strategically generated and combined into one composite learner. There are two features that are essential to an ensemble's performance, namely the individual accuracies of the component base learners and the overall diversity level of the ensemble. It is evident that diversity must have a role to play in the effectiveness of an ensemble, since multiple identical copies of the same learner clearly cannot improve on the performance of a single copy of that learner. However, the right balance of learner accuracy and ensemble diversity can improve the performance of machine learning tasks on benchmark data sets and real-world data sets, and recent theoretical and practical work has demonstrated the subtle trade-off between accuracy and diversity in an ensemble. Herein, we extend the extant literature by providing a deeper theoretical understanding for assessing and improving the optimality of any given ensemble, including random forests and deep neural network ensembles. We also propose a training algorithm for neural network ensembles and demonstrate that our approach provides improved performance when compared to both state-of-the-art individual learners as well as ensembles of state-of-the-art learners trained using standard loss functions. Our key insight is that it is better to explicitly encourage diversity in an ensemble, rather than merely allowing diversity to occur by happenstance, and that rigorous theoretical bounds on the trade-off between diversity and learner accuracy allow one to know when an optimal arrangement has been achieved. While our focus is on image classification problems, we wish to emphasize that the theory and approaches provided herein are also applicable to many other machine learning tasks where neural networks already provide state-of-the-art performance.