北京理工大學

数学与统计学院学术报告

Combining pre-trained large models via localized model averaging
通过局部模型平均组合预训练大模型

报告人： 高子文 (清华大学)

时间： 2026.3.9 下午16:00--17:00

地点：文萃楼D 703

摘要： Many pre-trained large models (PLMs) are being released. A popular approach to taking advantage of such PLMs is fine-tuning. However, when the data suitable for the task of interest are rather limited, fine-tuning may not be effective. In this case, weighting the predictions from different PLMs can be a better way to improve predictive performance. Motivated by such applications, we propose a localized model averaging method with weights modeled as functions of the covariates, making it substantially more versatile than existing model averaging methods. Under a general loss framework, we show that the flexible localized weights can be properly trained, and establish an asymptotic optimality property of the proposed method for both in-sample and out-of-sample risks. Extensive numerical experiments further demonstrate the effectiveness of the proposed method.

报告人简介： 高子文，清华大学博士后，博士毕业于中国科学院数学与系统科学研究院，研究方向是模型平均、**mixture of experts**、半监督学习等。目前主持国家自然科学基金 C 类项目、中国博士后科学基金面上资助项目，并在**IEEE Transactions on Reliability (**可靠性统计顶刊**)**、**The Lancet Microbe (**传染病学顶刊**)**、**Statistics in Medicine**等高水平期刊上发表多篇学术论文。

主办单位：北京理工大学数学与统计学院

School of Mathmatics and Statistics, Beijing Institute of Technology