Research Experience

Rong Wei

Email: rongwei2318@gmail.com

Last summer, I pursued an online research internship under the guidance of Prof. Jie Ding from UMN. Inspired by a new privacy framework, subset privacy, which originated from protecting participants' privacy in sensitive surveys, we investigated the properties of this privacy protection framework with rigorous mathematical definition and its possible applications in genomic data. Assuming that causal predictors are sparse and high-dimensional categorical data, we primarily intended to use the Lasso method, which could solve the regression problems and reduce the dimension of features. However, the literature review helped me change my previous assumptions and gave me more approaches to selecting the most significant features. Based on those findings, I gradually formulated vague ideas into a clear statistical regression problem with high-dimensional categorical predictors. On obfuscated genomic data, we investigated the performance of different baseline approaches. However, the model's performance was not satisfactory, and I came up with a hypothesis of why - that the uniform subset design might impair the significant relationship between predictors and the calculated response. With this assumption, I tried several different ways of subset design and achieved higher accuracy. Throughout my time in Prof Ding's group, I broadened my horizon of statistical application in genomics and became interested in AI security.