Project 1: Online change-point detection The detection of abrupt changes is a crucial open problem in almost every branch of science. In statistics, there is a well-developed research direction called the 'change-point problem'. However, the drawback of the traditional statistical methods is that they rely on a number of assumptions that are hard to validate in practical situations, for example, smoothness or distributional assumptions. This project aimed to detect the change-point by utilizing several machine learning techniques, for example, matrix factorization, and the problem will be solved under an online machine learning framework, where data becomes available in sequential order, in another word, we aimed to address the practical need for detecting change-point in a prospective way. In contrast, the majority of approaches in the existing literature are focused on an offline setting.

1. Aminikhanghahi, S., & Cook, D. J. (2017). A survey of methods for time series change point detection. Knowledge and information systems, 51(2), 339-367.
2. Chi, Y., Lu, Y. M., & Chen, Y. (2019). Nonconvex optimization meets low-rank matrix factorization: An overview. IEEE Transactions on Signal Processing, 67(20), 5239-5269.

Project 2: Dirty statistical models
This project aimed to develop a methodology to integrate data sets from several studies. However, the sample size of each data set is small so that statistical analysis for each data set along will lead to low statistical power. At the same time, each data set may be biased so a naive pooling procedure will not work. We want to develop new methods under survival analysis models (for example, Cox model, proportional odds model, linear transformation model, accelerated failure time model, etc.). The object of this project also includes the study of the non-asymptotic properties of the estimators.

1. Yang, E., & Ravikumar, P. (2013, December). Dirty statistical models. In Proceedings of the 26th International Conference on Neural Information Processing Systems-Volume 1 (pp. 611-619).
2. Chen, A., Owen, A. B., & Shi, M. (2015). Data enriched linear regression. Electronic journal of statistics, 9(1), 1078-1112.
3. Asiaee, A., Oymak, S., Coombes, K. R., & Banerjee, A. (2018). High Dimensional Data Enrichment: Interpretable, Fast, and Data-Efficient. arXiv preprint arXiv:1806.04047.
4. Wainwright, M. J. (2019). High-dimensional statistics: A non-asymptotic viewpoint (Vol. 48). Cambridge University Press.