Les mathématiques de l'apprentissage machine
« Atelier sur les défis modernes en théorie d'apprentissage »
23 au 26 avril 2018

Mathematics of Machine Learning
"Workshop on Modern Challenges of Learning Theory"
April 23-26, 2018

# Robustness in unsupervised and supervised machine learning

Gautam Kamath *

g@csail.mit.edu

---

Recently, the need for robust machine learning algorithms has become apparent. Whether due to errors in data collection, model misspecification, or adversarial attacks, contaminated datasets arise in many areas. This is an issue, as existing methods appear to be quite brittle to small amounts of errors. Even more worryingly, these models are being deployed in many security-critical settings, such as self-driving cars, where reliability is an absolute must.

In this talk, I will describe a line of work in which we provide provable guarantees for robust machine learning in several fundamental settings. I'll begin by discussing the problem of robust estimation of mean and covariance of a Gaussian distribution, and how to relax this to distributions with weaker assumptions on the moments. I will then describe how these methods can be used to "robustify" supervised learning algorithms by applying robust mean estimation algorithms to the gradients of the dataset. While theoretically sound, the algorithms are also realizable and efficient, and I will present experimental results on both synthetic and real-world data.

*Based on joint works with Ilias Diakonikolas, Daniel M. Kane, Jerry Li, Ankur Moitra, Jacob Steinhardt, and Alistair Stewart.*

*Computer Science & Artificial Intelligence Lab, MIT, 32 Vassar Street, Cambridge, MA 02139, USA