Algorithmic Aspects of Machine Learning: Final Project

Instructor: Ankur Moitra

Due: December 13th

The last assignment for the semester is to write a $4-6$ page final paper. You can work in pairs or you can work alone, and there are two options:

(1) You can write a literature review on some topic related to the material that we covered in class. In this class, we only focused on problems where there are *provable guarantees* so you should certainly choose a topic where there are provable guarantees. The goal of the writeup is to survey what is known, some main ideas of the proof in the papers you choose to read, and identify important open questions. You should think of this project as if you were going to give $1-2$ lectures in the course, and how you would explain some topic beyond what we covered in class, to other students. We covered a lot of material in class, but there is still much more out there and I hope to make your literature reviews available on the course website so that they can be a useful reference for others. Also, you should choose a somewhat focused topic because if you choose a topic that is too broad, it will be difficult to get into the details.

(2) You can do original research. This is the more challenging type of project, but it is also the most open ended. You should choose some open question connected to any of the topics we covered. You should try to give new *provable guarantees* for some problem, either by giving a new algorithm or giving an improved analysis of an existing algorithm. As with all research, it is never clear whether you will reach the goal that you set out. And so you should make sure that there is some partial progress that you can make along the way, that you can write up in your final paper. Even if you end up not proving what you set out to prove, you can still write up some preliminary ideas along the way or at the very least write up a literature review focused on the open question you worked on. If you choose this type of project, it is important that you come talk to either me or Alex about it so that we can discuss it with you and point you to other papers that might be relevant.

Please email me your final paper by Wednesday, December 13th. This is the last day of classes. Many of you have already come to talk to me about topics for your final project, and if you have ideas of your own of what you would like to work on, that is great! This is especially good because the purpose of the class was to expose you to some of the provable guarantees that are known for various problems in machine learning, and if you can find topics in your own research area where there are interesting avenues to explore, then you are more likely to use the material from this course in your own research even after the semester. That was one of my main goals in teaching this. If you do not have an idea, both Alex and I are here to help. Here are some suggested topics, for writing a literature review on. If you are interested in one of these topics, email me to claim it and set up a meeting so that we can talk about it. If you reserve a topic by yourself, but are interested in working with another student, I will put a '?' next to your name. And if we run out of topics, I will come up with more suggestions.

- **Spectral Clustering**
  "Improved Spectral Norm Bounds for Clustering", Awasthi, Sheffet.
  "Clustering with Spectral Norm and the k-Means Algorithm", Kumar, Kannan.

- **Clustering and Stability**
  "Approximate Clustering without the Approximation", Balcan, Blum, Gupta
  "Stability yields a PTAS for k-Median and k-Means", Awasthi, Blum, Sheffet.

- **Graph Partitioning and Stability**
  "Are Stable Instances Easy?", Bilu, Linial
  "Bilu-Linial Instances of MAX CUT and Minimum Multiway-cut", Makarychev, Makarychev, Vijayaraghavan.
  Claimed: Matthew Brennan

- **Smoothed Analysis in Learning**
  "Smoothed Analysis of Tensor Decompositions", Bhaskara, Charikar, Moitra, Vijayaraghavan
  "Learning Mixtures of Gaussians in High Dimensions", Ge, Huang, Kakade.

- **Graphical Models**
  "Reconstruction of Markov Random Fields from Samples: Some Observations and Algorithms", Bresler, Mossel, Sly.
  "Efficiently Learning Ising Models on Arbitrary Graphs", Bresler.

- **Isotonic Regression and Applications**
  " Efficient Learning of Generalized Linear and Single Index Models with Isotonic Regression", Kakade, Kalai, Kanade, Shamir.
  "Learning Graphical Models using Multiplicative Weights", Klivans, Meka.

- **Stochastic Block Model: Partial Recovery**
  "Stochastic Block Models and Reconstruction", Mossel, Neeman, Sly.
  Claimed: Paxton Turner

- **Stochastic Block Model: Exact Recovery**
  "Achieving Exact Cluster Recovery Threshold via Semidefinite Programming", Hajek, Wu, Xu.

- **Approximate Gradient Descent**
  "Phase Retrieval using Alternating Minimization", Netrapalli, Jain, Sanghavi.

"Statistical Guarantees for the EM Algorithm: From Population to Sample-based Analysis", Balakrishnan, Wainwright, Yu.
Claimed: Andrew Song

- **Escaping Saddle Points**
  "Gradient Descent Converges to Minimizers", Lee, Simchowitz, Jordan, Recht.
  "How to Escape Saddle Points Efficiently", Ge, Jin, Netrapalli, Kakade, Jordan.
  Claimed: Linus Hamilton, Jourdain Lamperski

- **Runtime/Statistical Tradeoffs**
  "SVM Optimization: Inverse Dependence on Training Set Size", Shalev-Shwartz, Srebro.
  "More Data Speeds up Training Time in Learning Halfspaces over Sparse Vectors", Daniely, Linial, Shalev-Shwartz.

- **Computational Lower Bounds**
  "Computational Lower Bounds for Sparse PCA", Berthet, Rigollet.
  "Lower Bounds on the Performance of Polynomial-time Algorithms for Sparse Linear Regression", Zhang, Wainwright, Jordan.

- **Sampling and Optimization**
  "Further and Stronger Analogy Between Sampling and Optimization: Langevin Monte Carlo and Gradient Descent", Dalalyan.
  "Beyond Log-concavity: Provable Guarantees for Sampling Multi-modal Distributions using Simulated Tempering Langevin Monte Carlo", Ge, Lee, Risteski
  Claimed: YounHun Kim

- **More Langevin Dynamics**
  "A Hitting Time Analysis of Stochastic Gradient Langevin Dynamics", Zhang, Liang, Charikar.
  "Non-convex Learning via Stochastic Gradient Langevin Dynamics: A Nonasymptotic Analysis", Raginsky, Rakhlin, Telgarsky.

- **Generalization in Deep Learning**
  "Understanding Deep Learning Requires Rethinking Generalization", Zhang, Bengio, Hardt, Recht, Vinyals.
  "Spectrally-normalized Margin Bounds for Neural Networks", Barlett, Foster, Telgarsky.

- **Provable Guarantees in Deep Learning**
  "Provable Bounds for Learning Some Deep Representations", Arora, Bhaskara, Ge, Ma.
  "Eigenvalue Decay Implies Polynomial-Time Learnability for Neural Networks", Goel, Klivans.