

Final Projects

Here are some ideas for final projects. I would like you to read one of the following papers (you may partner up) and write up a 4-6 page summary of what the paper is about and the main ideas. Additionally, you should set up a time to meet with me to talk about the paper after you've read it. Lastly, email me to claim one of these papers so that no paper is claimed more than once! And if you have another paper that you'd like to read instead, as long as it is relevant to the class material that is fine (just check with me).

- “Combining Labelled and Unlabelled Data with Co-Training”, Blum, Mitchell
- “Correlation Clustering”, Bansal, Blum, Chawla
- “Online Convex Programming and Generalized Infinitesimal Gradient Ascent”, Zinkevich
Claimed: Yin-Tat Lee
- “Efficient Algorithms for Online Decision Problems”, Kalai, Vempala
Claimed: Vlad Firoiu, Nathan Pinsker
- “Approximate Clustering without the Approximation”, Balcan, Blum, Gupta
- “A Discriminative Framework for Clustering via Similarity Functions”, Balcan, Blum, Vempala
- “On Clusterings: Good, Bad and Spectral”, Kannan, Vempala, Vetta
- “Clustering Large Graphs via the Singular Value Decomposition”, Drineas et al
“NP-hardness of Euclidean Sum-of-Squares Clustering” Aloise et al
Claimed: Matt Coudron, Adrian Vladu
- “Latent Semantic Indexing: A Probabilistic Analysis”, Papadimitriou et al
Claimed: Will Grathwohl
- “Using Mixture Models for Collaborative Filtering”, Kleinberg, Sandler
Claimed: Jing Lin

- “On Learning Mixtures of Heavy-Tailed Distributions”, Dasgupta, Hopcroft, Kleinberg, Sandler
Claimed: Gautam Kamath
- “The Spectral Method for General Mixture Models”, Kannan et al
- “Spectral Partitioning of Random Graphs”, McSherry
“The Eigenvalues of Random Symmetric Matrices”, Furedi, Komlos
Claimed: Yufei Zhao
- “Spectral Analysis of Data”, Azar et al
- “Random Projection Trees and Low-Dimensional Manifolds”, Dasgupta, Freund
Claimed: Josh Alman, Timothy Chu
- “SVM Optimization: Inverse Dependence on Training Set Size”, Shalev-Shwartz, Srebro
- “Learning Markov Networks: Maximum Bounded Tree-Width Graphs”, Karger, Srebro
Claimed: Adam Hesterberg, David Rolnick
- “Rank, Trace-Norm and Max-Norm”, Shraibman, Srebro
Claimed: Adam Bouland, Henry Yuen
- “Matrix Completion from a Few Entries”, Keshavan, Montanari, Oh
- “Near-Optimal Column-Based Matrix Reconstruction”, Drineas, Boutsidis, and Malik Magdon-Ismail
Claimed: Cameron Musco, Christopher Musco
- “Learning Intersections and Thresholds of Halfspaces”, Klivans, O’Donnell, Servedio
Claimed: Madalina Persu
- “Learning Geometric Concepts via Gaussian Surface Area”, Klivans, O’Donnell, Servedio
- “Restriction Access”, Dvir, Rao, Wigderson, Yehudayoff
- “Identifiability and Unmixing of Latent Parse Trees”, Hsu, Kakade, Liang

- “Latent Variable Graphical Model Selection via Convex Optimization”, Chandrasekaran, Parrilo, Willsky
- “Computational and Statistical Tradeoffs via Convex Relaxation”, Chandrasekaran, Jordan
Claimed: Jan-Christian Hutter
- “Reconstruction of Markov Random Fields from Samples: Some Observations and Algorithms”, Bressler, Mossel, Sly
- “Provable Bounds for Learning Some Deep Representations”, Arora, Bhaskara, Ge, Ma
- “Tensor Decompositions for Learning Latent Variable Models.”, Anandkumar, Ge, Hsu, Kakade, Telgarsky
- “Complexity of Inference in Latent Dirichlet Allocation”, Sontag, Roy
Claimed: Pratiksha Thaker
- “Unsupervised Learning of Noisy-Or Bayesian Networks”, Halpern, Sontag
- “Multi-view Clustering via Canonical Correlation Analysis”, Chaudhuri, Kakade, Livescu, Sridharan