Soumyadeep Chatterjee

CONTACT INFORMATION
718 Old San Francisco Road, Apt 273
Sunnyvale, CA 94086
soumyachat@gmail.com
Phone: 612-321-1212

SUMMARY

I have 7+ experience in developing Recommendation Systems and Machine Learning. I develop large scale pipelines for data processing, model training and serving. I have published 8+ papers in top conferences in Machine Learning and Data Mining.

EMPLOYMENT

Machine Learning Engineer Jan 2018 – present
Quora
- Continuously trained Learning to Rank (LTR) system for Quora's Ask to Answer (A2A) product, increasing key metrics by over 70%
- Incorporating state of the art content embeddings (e.g. BERT) in ranking systems
- Developing ranking models for 10 new languages for Quora's Internationalization team

Data Scientist Jul 2016 – Dec 2017
Quora
- Developed first generation of ML models for spam detection, reducing spam on Quora product by more than 10%
- Developed Markov models for user activity on Quora using user follows and actions, to improve follow suggestions and content distribution.
- Managed Data Research program that conducted research to develop insights for long-term company strategy

Research Scientist Oct 2015 – Jul 2016
Personalization Sciences, Yahoo!
- Developed machine learning models for serving image and news results in Yahoo! mobile search improving metrics by 50%
- Developed end to end pipeline in Hadoop for extracting data from user activity logs, training ranking model and deployment

Research Assistant May 2010 – Aug 2015
Dept. of Computer Science & Engg., University of Minnesota, Twin Cities
- Statistical methods for regularization in high dimensional regression problems, applications to improving climate predictions

EDUCATION

University of Minnesota, Minneapolis, MN Sep 2009 – Sep 2015

Ph.D. in Computer Science & Engineering
- Thesis: High Dimensional Statistical Models: Applications to Climate

Jadavpur University, Kolkata, India Jul 2005 – Jun 2009

B. E. in Electronics & Telecommunications Engineering

Skills

- **Languages:** Python, C++, MySql, Hive, MATLAB, Java (basic), Pig (basic)
- **Frameworks:** Hadoop, Spark, Tensorflow, Hbase, Kafka
- **Operating Systems:** Linux, Windows, Mac OS X

Selected Publications

1. **S. Chatterjee**, V. Sivakumar, Andre R. Gon calves and A. Banerjee, "Structured Estimation in High Dimensions and Multitask Learning with Applications in Climate", Large-Scale Machine Learning in the Earth Sciences. Chapman & Hall/CRC, 2016.
2. **S. Chatterjee**, S. Liess, A. Banerjee and V. Kumar, "Understanding Dominant Factors for Precipitation over the Great Lakes Region", Thirtieth AAAI Conference On Artificial Intelligence 2016.
3. A. Asiaee Taheri, **S. Chatterjee** and A. Banerjee, "Regularized Structured Estimation in High-Dimensions with Noisy Designs", SIAM Intl. Conference on Data Mining 2016.
4. **S. Chatterjee**, S. Chen and A. Banerjee, "Generalized Dantzig Selector: Application to the k-support norm", Advances in Neural Information Processing Systems (NIPS), 2014.
5. A.Goncalves, P. Das, **S. Chatterjee**, V. Sivakumar, F. J. Von Zuben and A. Banerjee, "Multi-task Sparse Structure Learning", International Conference on Information and Knowledge Management (CIKM), 2014.
6. H. Wang, F. Fazayeli, **S. Chatterjee** and A. Banerjee, "Gaussian Copula Precision Estimation with Missing Values", International Conference on Artificial Intelligence and Statistics (AISTATS), 2014.
7. **S. Chatterjee**, A. Banerjee, S. Chatterjee and A. Ganguly, "Mixture of Regression Models for Precipitation Prediction", The Second International Workshop on Climate Informatics (CI), 2012.
8. **S. Chatterjee**, K. Steinhaeuser, A. Banerjee, S. Chatterjee and A. Ganguly, "Sparse Group Lasso: Consistency and Climate Applications", SIAM Intl. Conference on Data Mining 2012 (**Best Student Paper Award**).

Awards & Scholarships

- Best Student Paper Award at the SIAM International Conference on Data Mining (SDM), 2012.
- 3 year Graduate School Fellowship (2009-2012) from Univ. of Minnesota Twin Cities.
- Travel Award for paper presented at SDM 2012.
- 4-year J. C. Bose National Science Talent Search (JBNSTS) Senior Scholarship, 2005–2009.