Sergey Kirshner

Résumé

Santa Clara, CA
email.me@sergeykirshner.com
sergeykirshner.com
sergeykirshner

Summary

I am a statistical machine learning expert with 17+ years of experience (12+ years post-Ph.D.) in modeling and understanding of large-scale noisy data in high-impact applications. My previous work spans applications of probabilistic approaches (including graphical and time series models) and Bayesian techniques to high-dimensional data understanding, with applications including relevance modeling for e-commerce and ad tech, rainfall and atmospheric data modeling, understanding of statistical properties of graph/networks.

Education

University of California, Irvine
Ph.D., Information and Computer Science *2005*

University of California, Irvine
M.S., Information and Computer Science *2001*

University of California, Berkeley
B.A., Mathematics & Computer Science (double) *1998*

Recent Professional Experience

Facebook **Menlo Park, CA, USA**
Research Scientist *2017 – present*

@WalmartLabs **Sunnyvale, CA, USA**
Director of Modeling/Principal Architect *2016 – 2017*
Driving modeling and analytics projects for the Walmart Advertisement Platform for partnership marketing, including measurement, optimized audience segment construction, and real-time bidding on demand side platforms; building a team of scientists and engineers to convert Walmart's transaction and other data into ad tech products generating revenue for the company.

Skytree (acquired by Infosys) **San Jose, CA, USA**
Principal Member of Technical Staff *2014 – 2015*
Research and prototyping of machine learning approaches for classification and anomaly detection and for automation of the data science process including feature generation and selection, hyperparameter selection, and model estimation; participation (including lead) in customer projects and POCs; development of IP.

a9.com (Amazon) **Palo Alto, CA, USA**
SDE/Machine Learning Scientist *2013 – 2014*
Product search ranking by relevance: development and implementation of ranking approaches in a massive data setting with strict latency requirements, A/B hypothesis testing, deployment of ranking functions into production with impact to millions of customers daily, communication with business teams for decision making and planning.

Purdue University, Statistics **West Lafayette, IN, USA**
Assistant Professor *2008 – 2013*
Research: original statistical machine learning research with applications to modeling of multivariate atmospheric data and to understanding of large scale network behavior; funded collaborations with scientists from other disciplines.
Training: mentorship and supervision of graduate and undergraduate students, teaching and development of graduate and undergraduate courses in statistics and machine learning.

Selected Publications and Patents

Full list (20 peer-reviewed conference and journal publication; patents; workshop papers and tech reports) is available at sergeykirshner.com/publications.

- **S. Kirshner,** A. Gray, L. Kite, 'Modeling of geospatial location over time', U.S. Patent Application №15/254,958, filed September 1, 2016.
- **S. Kirshner,** 'Constructing additive trees monotonic in selected sets of variables', U.S. Patent Application №15/178,549, filed June 9, 2016.
- S. Moreno, J. Neville, **S. Kirshner,** 'Learning mixed Kronecker product graph models with simulated method of moments', in *Proceedings of the Nineteenth ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 1052-1060, August 2013 (KDD-2013, accepted/submitted 126/726)
- G. Mallya, S. Tripathi, **S. Kirshner,** R.S. Govindaraju, 'Probabilistic Assessment of Drought Characterization using a Hidden Markov Model', *Journal of Hydrologic Engineering,* volume 18, pp. 834-845, July 2013. doi:10.1061/(ASCE)HE.1943-5584.0000699
- B. Póczos, **S. Kirshner**, Cs. Szepesvári, 'REGO: Rank-based estimation of Rényi information using Euclidean graph optimization', in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, pp. 605-612, Y. W. Teh and M. Titterington (editors), July 2010, Journal of Machine Learning Research Workshop and Conference Proceedings, volume 9 (AISTATS-2010) (plenary session, plenary/accepted/submitted 24/125/308)
- **S. Kirshner**, B. Póczos, 'ICA and ISA using Schweizer-Wolff dependence measure', in *Proceedings of the Twenty-Fifth International Conference on Machine Learning*, pp. 464-471, A. McCallum and S. Roweis (eds.), July 2008 (ICML-2008) (accepted/submitted 155/583)
- **S. Kirshner,** 'Learning with tree-averaged densities and distributions', in *Advances in Neural Information Processing Systems 20*, pp. 761-768, J.C. Platt and D. Koller and Y. Singer and S. Roweis (eds.), MIT Press, Cambridge, MA, 2008 (NIPS-2007) (plenary session, plenary/accepted/submitted 26/217/975)
- A.W. Robertson, **S. Kirshner,** P. Smyth, S.P. Charles, and B.C. Bates, 'Subseasonal-to-interdecadal variability of the Australian monsoon over North Queensland.' *The Quarterly Journal of Royal Meteorological Society*, volume 132, number 615, pp. 519-542, January 2006. doi:10.1256/qj.05.75
- **S. Kirshner,** P. Smyth, A.W. Robertson, 'Conditional Chow-Liu tree structures for modeling discrete-valued vector time series,' in *Proceedings of the Twentieth Conference on Uncertainty in Artificial Intelligence*, pp. 317-324, M. Chickering, J. Halpern (eds.), AUAI Press, July 2004 (UAI-2004) (plenary session, plenary/accepted/submitted 27/75/253)
- A.W. Robertson, **S. Kirshner,** and P. Smyth, 'Daily rainfall occurrence over Northeast Brazil and its downscalability using a hidden Markov model.' *Journal of Climate*, volume 17, issue 22, pp. 4407-4424, November 2004. doi:10.1175/JCLI-3216.1

Software

LTC: Matlab latent tree copula toolbox. *2012*

SWICA: Matlab independent component analysis copula-based methods toolbox. *2008-2012*

MVNHMM: C/C++ toolbox for modeling multivariate time-series with hidden Markov models. *2003-2008*

Technical Skills

Programming Languages: Python, C/C++

Computing/Statistical Software: NumPy, SciPy, pandas, scikit-learn, R, Matlab

Data Manipulation: Hive/SQL, Spark

Document Preparation: LaTeX