Sajid M. Siddiqi

Carnegie Mellon University
Robotics Institute
NSH 3122
Pittsburgh, PA 15217
Date of Birth: January 10, 1981
Immigration Status: US Permanent Resident

Cell: (412) 680-9963
Email: siddiqi@cs.cmu.edu
Homepage: http://www.cs.cmu.edu/~siddiqi

Objective

To seek a postdoctoral or permanent research position with a focus on challenging projects in statistical machine learning and data mining.

Research Summary

I am interested in modeling the behavior of time-evolving systems using latent-variable representations. My research focuses on circumventing drawbacks of traditional latent variable models: local minima and instability during optimization, difficulties in scaling to large state spaces or high dimensional data, etc. I create novel learning and inference algorithms for existing models and devise new models with theoretically sound learning methods. The resulting models are useful for tasks such as prediction, recognition, classification, anomaly detection, inferring missing values, and decision-making under uncertainty. Examples of systems I have worked on include mobile robot sensors, laptop user modeling, author-keyword evolution in document corpora, biosurveillance for disease tracking, dynamic video textures, and human activity recognition over time using wearable sensors.

Research Interests

Statistical Machine Learning, Sequential Data Modeling, Probabilistic Mobile Robotics, Activity Recognition, Unsupervised Learning, Manifold Learning, Spectral Methods.

Education

Ph.D. Robotics and Artificial Intelligence, Carnegie Mellon University, *expected* October 2009.

Dissertation: Latent Variable and Predictive Models of Dynamical Systems.

Committee: Geoffrey J. Gordon *(Chair)*, Andrew W. Moore, Jeff Schneider, Zoubin Ghahramani and David Wingate.

M.S. Robotics and Artificial Intelligence, Carnegie Mellon University, 2006.

GPA: 3.93

B.S. Computer Science, University of Southern California, 2003.

B.A. Mathematics and Economics, University of Southern California, 2003.
GPA: 3.88

Academic Experience

Carnegie Mellon University, Robotics Institute

Research Assistant to Geoffrey J. Gordon, Spring 2006–Present.

Teaching Assistant to Martial Hebert for Artificial Intelligence, Spring 2006.

Research Assistant to Andrew W. Moore, Fall 2003–Fall 2005.

University of Southern California, Department of Computer Science

Undergraduate Research Assistant to Gaurav Sukhatme, Spring 2001–Spring 2003.

Research

In Submission

Sajid M. Siddiqi, Byron Boots and Geoffrey J. Gordon. Reduced-Rank Hidden Markov Models.

Refereed Publications

Sajid M. Siddiqi, Byron Boots and Geoffrey J. Gordon (2008). A Constraint Generation Approach to Learning Stable Linear Dynamical Systems. In *Advances in Neural Information Processing Systems (NIPS) 21*.

Sajid M. Siddiqi, Geoffrey J. Gordon and Andrew W. Moore (2007). Fast State Discovery for HMM Model Selection and Learning. In *Proceedings of the 11th International Conference on Artificial Intelligence and Statistics (AI-STATS)*.

Purnamrita Sarkar, Sajid M. Siddiqi and Geoffrey J. Gordon (2007). A Latent Space Approach to Dynamic Embedding of Co-occurrence Data. In *Proceedings of the 11th International Conference on Artificial Intelligence and Statistics (AI-STATS)*.

Georgios Theocharous, Shie Mannor, Nilesh Shah, Prashant Gandhi, Branislav Kveton, Sajid M. Siddiqi, and Chih-Han Yu (2006). Machine Learning for Adaptive Power Management. *Intel Technology Journal*.

Sajid M. Siddiqi and Andrew W. Moore (2005). Fast Inference and Learning in Large-State-Space HMMs. In *Proceedings of the 22nd International Conference on Machine Learning (ICML)*.

Andrew Howard, Sajid M. Siddiqi and Gaurav Sukhatme (2003). An Experimental Study of Localization Using Wireless Ethernet. In *Proceedings of the 4th International Conference on Field and Service Robotics (FSR)*.

Sajid M. Siddiqi, Andrew Howard and Gaurav Sukhatme (2003). Experiments in Monte-Carlo Localization using WiFi Signal Strength. In *Proceedings of the 11th International Conference on Advanced Robotics (ICAR)*.

Other Publications

Sajid M. Siddiqi, Byron Boots and Geoffrey J. Gordon (2008). A Constraint Generation Approach to Learning Stable Linear Dynamical Systems. *CMU Technical Report CMU-ML-08-101*.

Sajid M. Siddiqi, Byron Boots, Geoffrey J. Gordon and Artur Dubrawski (2007). Learning Stable Multivariate Baseline Models for Outbreak Detection. In *Advances in Disease Surveillance (SDS) Volume 4*.

Purnamrita Sarkar, Sajid M. Siddiqi and Geoffrey J. Gordon (2006). Approximate Kalman Filters for Embedding Author-Word Co-occurrence Data over Time. In *Proceedings of the Workshop on Statistical Network Analysis: Models, Issues and New Directions at ICML 2006*.

Brigham Anderson, Sajid M. Siddiqi and Andrew W. Moore (2006). Sequence Selection for Active Learning. *CMU Technical Report CMU-RI-06-16*.

Industry Experience

Scientific Advisor, *MobileFusion Inc.*, Summer 2008-present (part-time)

I am supervising (a) the application of my algorithms and code to audio-based event detection, and (b) integration of this code into MobileFusion's software for their portable sensor device.

Research Intern, *Google Research*, Summer 2007

Used manifold-learning and sequential data modeling to learn classifiers for high-dimensional sequences such as videos and mp3s.

Mentor: Sanjiv Kumar

Research Intern, *Intel Research*, Summer 2006

Created and applied machine learning algorithms for modeling laptop user behavior to perform adaptive power management and extend battery life.

Mentor: Georgios Theocharous

Software Engineering Intern, *Microsoft Corp.*, Summer 2002

Web Programming Intern, *Synopsys Inc.*, Summer 2001

Software Engineering Intern, *Oblix Inc.*, Summer 2000

Talks

MobileFusion, Inc., April 2009
Modeling Sequential Data

NASA Jet Propulsion Laboratories, February 2009
Learning Latent Variable Models of Dynamical Systems

Carnegie Mellon University Machine Learning Lunch, April 2008
Learning Stable Linear Dynamical Systems

Conference on Neural Information and Processing Systems (NIPS), December 2007
A Constraint Generation Approach to Learning Stable Linear Dynamical Systems

Conference on Artificial Intelligence and Statistics (AISTATS), March 2007
Fast State Discovery for HMM Model Selection and Learning

Google Research NY, August 2007
Data Sequence Clustering and Classification

Intel Research Santa Clara, September 2006
Discovering Laptop User Context for Adaptive Power Management

Carnegie Mellon University Machine Learning Lunch, November 2005
Fast Inference and Learning in Large-State-Space HMMs

International Conference on Advanced Robotics (ICAR), June 2003
Experiments in Monte-Carlo Localization using WiFi Signal Strength

Honors

USC Presidential Scholarship (half-tuition), 1999-2003

First Place (Physical Sciences, Mathematics and Engineering category) at Annual USC Undergraduate Research Symposium, 2003

Outstanding Computer Science Graduate Award (USC), 2003

Robotics Graduate Fellowship (CMU) 2003-present

Professional Activities

Reviewer, *Journal of Machine Learning Research (JMLR)* 2008,2009

Program Committee member, *International Conference on Machine Learning (ICML)* 2007

Reviewer, *North-Eastern Student Colloquium on Artificial Intelligence (NESCAI)* 2006

Leadership

Secretary, *Association for India's Development (AID) Pittsburgh Chapter*, 2005-2006

President, *Robotics Graduate Student Organization (RoboOrg)*, 2004-2005

Vice-President, *Robotics Graduate Student Organization (RoboOrg)*, 2003-2004

Miscellaneous

Computer Skills: C, C++, Matlab, Perl, R, SQL, Linux, Windows

Courses: Machine Learning, Artificial Intelligence, Intermediate Statistics, Statistical Foundations of Machine Learning, Probabilistic Graphical Models, Computer Vision, Mathematics for Robotics, Graduate Algorithms, Machine Learning Theory

References

Available upon request.

Last updated: August 19, 2009