Liu Yang

Contact Information

7223 Gates-Hillman Center, Computer Science Department, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213-3891
Tel: 517-526-2509 (cell) (412) 268-7669 (office)
E-mail: liuy@cs.cmu.edu URL: http://www.cs.cmu.edu/~liuy

Research Interests

My area of research is Theoretical Computer Science. I am particularly interested in

- Statistical Learning Theory: Active Learning, Transfer Learning, Online Learning
- Property Testing
- Computational Learning Theory, with emphasis on Interactive Protocols
- Algorithmic Economics, Online Algorithms, Pricing Problems
- Optimization, Semi-supervised Learning with applications to Text Categorization, Distance Metric Learning (algorithms and applications to Computer Vision tasks such as Object Recognition and Biomedical Imaging and Analysis, focusing on Interactive Search-Assisted Diagnosis).

Current Position

- **Carnegie Mellon University, Pittsburgh, PA**
  Postdoctoral Fellow in the Computer Science Department.
  Supervisor: Avrim Blum

Education

- **Carnegie Mellon University, Pittsburgh, PA**
  Ph.D. in Machine Learning from the School of Computer Science, December 2013.
  Advisors: Avrim Blum and Jaime Carbonell
  Dissertation: Mathematical Theories of Interaction with Oracles
  Thesis Committee: Avrim Blum, Manuel Blum, Jaime Carbonell, Sanjoy Dasgupta, Yishay Mansour, and Joel Spencer.
- **Carnegie Mellon University, Pittsburgh, PA**
  Masters Degree in Machine Learning from the School of Computer Science, May 2010.
- **Huazhong University of Science and Technology, China**
  Bachelor of Engineering in Electronics Engineering. I was enrolled in the **Advanced Class** of Huazhong Univ. of Sci. and Tech. It consists of the top 3% students selected from its 6 depts.

References

- **Avrim Blum**
  Computer Science Department, Carnegie Mellon University
  5000 Forbes Avenue, Pittsburgh, PA 15213-3891
  E-mail: avrim@cs.cmu.edu
  Tel: (412) 268-6452
- **Jaime Carbonell**
  Language Technologies Institute, Carnegie Mellon University
  5000 Forbes Avenue, Pittsburgh, PA 15213-3891
  E-mail: jgc@cs.cmu.edu
  Tel: (412) 268-7279
- **Yishay Mansour**
  School of Computer Science, Tel Aviv University
  Ramat - Aviv, Tel - Aviv 69978, Tel - Aviv, Israel
  E-mail: mansour@tau.ac.il
  Tel: +972-3-6408829 (office)

Teaching

- Invited Guest Lectures on **Online Pricing, Correlated Auction** in 15-896 Algorithms, Games, and Networks (Spring 2013), taught by Ariel Procaccia and Avrim Blum, CS Dept., CMU.
- TA for 15-750 Graduate Algorithms (Spring 2011), taught by Manuel Blum, CS Dept., CMU. Invited Guest Lecture on **Online Learning Mistake Bound Model**.

- Three Invited Guest Lectures on **Property Testing, Active Learning, and Transfer Learning** in 15-859(B) Machine Learning Theory (Spring 2012), taught by Avrim Blum, CS Dept., CMU.
- TA for 15-355 Modern Computer Algebra (Fall 2011). Invited Guest Lecture on **Gröbner Bases**.

Papers Under Review

- **Online Allocation and Pricing with Economies of Scale**. With Avrim Blum and Yishay Mansour. In submission to The 46th Annual ACM Symposium on Theory of Computing (STOC), 2014.
- **Surrogate Losses in Passive and Active Learning**. With Steve Hanneke. In submission to The Annals of Statistics. [arXiv:1207.3772]
- **Learning with a Drifting Target Concept**. With Steve Hanneke and Varun Kanade. In submission to the Journal of Machine Learning Research.
- **Minimax Analysis of Active Learning**. With Steve Hanneke. In submission to the Journal of Machine Learning Research.
- **Estimation of Priors with Applications to Preference Elicitation**. With Steve Hanneke. In submission to the 28th AAAI Conference on Artificial Intelligence.

Recent Publications

- **Buy-in-Bulk Active Learning**. Liu Yang and Jaime Carbonell. Advances in Neural Information Processing Systems 26 (NIPS), 2013.
- **Learnability of DNF with Representation-Specific Queries**. With Avrim Blum and Jaime Carbonell. The 4th Innovations in Theoretical Computer Science (ITCS), 2013.
- **Activized Learning with Uniform Classification Noise**. Liu Yang and Steve Hanneke. The 30th International Conference on Machine Learning (ICML), 2013.
- **Active Property Testing**. With Nina Balcan, Eric Blais, and Avrim Blum. The 53rd Annual Symposium on Foundations of Computer Science (FOCS), 2012. [arXiv:1111.0897]
- **A Theory of Transfer Learning with Applications to Active Learning**. Liu Yang, Steve Hanneke, and Jaime Carbonell. Machine Learning Journal, 2012.
- **Bounds on the Minimax Rate for Estimating a Prior over a VC Class from Independent Learning Tasks**. Liu Yang, Steve Hanneke, and Jaime Carbonell. Tech Report CMU-ML-12-112.
- **Identifiability of Priors from Bounded Sample Sizes with Applications to Transfer Learning**. Liu Yang, Steve Hanneke, and Jaime Carbonell. The 24th Annual Conference on Learning Theory (COLT), Budapest, Hungary, 2011.
- **Active Learning with a Drifting Distribution**. Liu Yang. Advances in Neural Information Processing Systems 24 (NIPS), 2011.
- **The Sample Complexity of Self-Verifying Bayesian Active Learning**. Liu Yang, Steve Hanneke, and Jaime Carbonell. The 14th International Conference on Artificial Intelligence and Statisitcs (AISTATS), 2011.
- **Bayesian Active Learning Using Arbitrary Binary Valued Queries**. Liu Yang, Steve Hanneke, and Jaime Carbonell. Proceedings of the 21st International Conference on Algorithmic Learning Theory (ALT), 2010.
- **Negative Results for Active Learning with Convex Losses**. With Steve Hanneke. The 13th International Conference on Artificial Intelligence and Statisitcs (AISTATS), 2010.
- **A Boosting Framework for Visuality-Preserving Distance Metric Learning and its Application to Medical Image Retrieval**. Liu Yang, Rong Jin, Lilly Mummert, Rahul Sukthankar, Adam Goode, Bin Zheng, Stephen Hoi and Mahadev Satyanarayanan. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010 January, 32(1):30-44.

- **Online Learning by Ellipsoid Method**. Liu Yang, Rong Jin, and Jieping Ye. The 26th International Conference on Machine Learning (ICML), 2009.
- **Cost Complexity of Proactive Learning via a Reduction to Realizable Active Learning**. Liu Yang and Jamie Carbonell. Tech Report CMU-ML-09-113.
- **Adaptive Proactive Learning with Cost-Reliability Tradeoff**. Liu Yang and Jamie Carbonell. Tech Report CMU-ML-09-114.

Selected Past Publications: Algorithm Design and Optimization Learning & Vision

- **Semi-supervised Learning with Weakly-Related Unlabeled Data: Towards Better Text Categorization**. Liu Yang, Rong Jin and Rahul Sukthankar. Advances in Neural Information Processing Systems 21 (NIPS), 2008.
- **Unifying Discriminative Visual Codebook Generation with Classifier Training for Object Category Recognition**. Liu Yang, Rong Jin, Rahul Sukthankar and Frederic Jurie. Proceedings of Computer Vision and Pattern Recognition (CVPR), 2008. (Oral Presentation)
- **Bayesian Active Distance Metric Learning**. (Oral Presentation). Liu Yang, Rong Jin and Rahul Sukthankar. Proceedings of the 23rd Conference on Uncertainty in Artificial Intelligence (UAI), 2007.
- **Discriminative Cluster Refinement: Improving Object Category Recognition Given Limited Training Data**. Liu Yang, Rong Jin, Caroline Pantofaru, Rahul Sukthankar. Proceedings of Computer Vision and Pattern Recognition (CVPR), 2007.
- **Learning Distance Metrics for Interactive Search-assisted Diagnosis of Mammograms**. Liu Yang, Rong Jin, Rahul Sukthankar, Bin Zheng, Lily Mummert, Mahadev Satyanarayanan, Mei Chen, and Drazen Jukic. Conference on Computer-Aided Diagnosis, SPIE Symposium on Medical Imaging, 2007.
- **Resource-constrained supervised dimensionality reduction**. (Oral Presentation). Liu Yang, Rong Jin, Rahul Sukthankar. The First International Workshop on Multimodal Information Retrieval at IJCAI, 2007.
- **An Efficient Algorithm for Local Distance Metric Learning**. (Oral Presentation). Liu Yang, Rong Jin, Rahul Sukthankar, Yi Liu. The 21st National Conference on Artificial Intelligence (AAAI), 2006.
- **Semi-supervised Multi-label Learning by Constrained Non-negative Matrix Factorization**. (Oral Presentation). Yi Liu, Rong Jin, Liu Yang. The 21st National Conference on Artificial Intelligence (AAAI), 2006.
- **Algorithm of Image Registration Based on Edge Matching and Multi-scale Wavelet Transformation**. Liu Yang, Furong Wang, Benxiong Huang, November, 2004. Journal of Huazhong University of Science and Technology (Natural Science Edition).

Seminar Talks

- **Online Allocation and Pricing with Economies of Scale** **November 14, 2013**
  Algorithms, Combinatorics, and Optimization (ACO) Seminar, Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, Pennsylvania.
- **Active Testing Real-valued Functions** **October 30, 2013**
  CMU Theory Lunch, Algorithms and Complexity Theory Group, Computer Science Department, Carnegie Mellon University, Pittsburgh, Pennsylvania.
- **Active Testing Boolean and Real-valued Functions** **October 21, 2013**
  CS Theory/Math Seminar, Department of Computer Science, Purdue University, West Lafayette, Indiana.

- **Online Allocation and Pricing with Economies of Scale** **September 13, 2013**
  Economics and Computer Science Research Seminar, EconCS Group, School of Engineering and Applied Science, Harvard University, Cambridge, Massachusetts.
- **Combinatorial Approaches to Active Learning and Transfer Learning** **February 19, 2013**
  Machine Learning Ph.D. Seminar, the Courant Institute of Mathematical Sciences, New York University, New York City, New York.
- **Mathematical Theories of Interaction with Oracles: Active Property Testing and New Models for Learning Boolean Functions** **February 11, 2013**
  Computer Science/Discrete Mathematics Seminar, School of Mathematics, Institute for Advanced Study, Princeton, New Jersey.
- **Active Property Testing** **January 14, 2013**
  CS Theory Seminar, IBM Almaden Research Center, Almaden, California.
- **Active Property Testing** **October 4, 2012**
  Algorithms and Complexity Seminar, MIT CSAIL Theory of Computation, Massachusetts Institute of Technology, Cambridge, Massachusetts.
- **Active Property Testing** **September 24, 2012**
  Berkeley Theory Seminar, EECS Department, University of California at Berkeley, Berkeley, California.
- **Active Learning, Drifting Distributions, and Convex Losses** **April 30, 2012**
  CMU Machine Learning Lunch Seminar, Machine Learning Department, Carnegie Mellon University, Pittsburgh, Pennsylvania.
- **Learnability of DNF with Representation-Specific Queries** **April 11, 2012**
  CMU Theory Lunch, Algorithms and Complexity Theory Group, Computer Science Department, Carnegie Mellon University, Pittsburgh, Pennsylvania.
- **Active Testing** **November 9, 2011**
  CMU Theory Lunch, Algorithms and Complexity Theory Group, Computer Science Department, Carnegie Mellon University, Pittsburgh, Pennsylvania.
- **Identifiability of Priors from Bounded Sample Sizes with Applications to Transfer Learning** **October 31, 2011**
  CMU Machine Learning Lunch Seminar, Machine Learning Department, Carnegie Mellon University, Pittsburgh, Pennsylvania.
- **Discriminative Cluster Refinement: Improving Object Category Recognition Given Limited Training Data** **July 19, 2007**
  Vision and Media Lab, Simon Fraser University, Burnaby, BC, Canada.
- **Improving Object Recognition by Discriminative Cluster Refinement** **June 11, 2007**
  Vision and Autonomous Systems Center (VASC) Seminar, The Robotics Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania.
- **Local Distance Metric Learning** **July 21, 2006**
  Biomedical Imaging and Analysis Seminar, Computer Science and Artificial Intelligence Laboratory (CSAIL), Massachusetts Institute of Technology, Cambridge, Massachusetts.
- **Applying Local Distance Metric Learning to Image Retrieval** **July 21, 2006**
  Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts.

Research Group Talks

- **A Theory of Transfer Learning with Applications to Active Learning** **February 27, 2013**
  Center for Biological & Computational Learning, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology.

- **Active Learning, Drifting Distributions, and Convex Losses** **January 23, 2013**
  Peter Bartlett's Research Group, Computer Science Division, University of California at Berkeley, Berkeley, California.

Conference Talks

- **Mathematical Theories of Interaction with (Human) Oracles** **January 9, 2012**
  Graduating bits: Finishing Ph.D.'s and Postdoc Short Presentation in the 3rd Innovations in Theoretical Computer Science (ITCS) conference, Cambridge, Massachusetts.
- **Online Learning by Ellipsoid Method** **June 15, 2009**
  Oral Presentation in the 26th International Conference on Machine Learning, Montreal, Canada.
- **Bayesian Active Distance Metric Learning** **July 20, 2007**
  Oral Presentation in the 23rd Conference on Uncertainty in Artificial Intelligence, University of British Columbia Vancouver, BC, Canada.
- **An Efficient Algorithm for Local Distance Metric Learning** **July 18, 2006**
  AAAI-06: Twenty-First National Conference on Artificial Intelligence, Boston, Massachusetts.

Program Committee

ICML 2012 (The 29th International Conference on Machine Learning), ICML 2013 (The 30th International Conference on Machine Learning).

Journal Refereeing

Journal of Machine Learning Research, IEEE Transactions on Pattern Analysis and Machine Intelligence, Machine Vision and Applications, Pattern Recognition, Signal Processing, International Journal of Pattern Recognition and Artificial Intelligence, IEEE Transactions on Neural Networks.

Conference Refereeing

ICML 2009 (The 26th International Conference on Machine Learning), COLT 2011 (The 24th Annual Conference on Learning Theory), CVPR 2012 (The 25th IEEE Conference on Computer Vision and Pattern Recognition), CVPR 2013 (The 26th IEEE Conference on Computer Vision and Pattern Recognition), ICCV 2013 (IEEE International Conference on Computer Vision), SODA 2013 (SIAM: ACM-SIAM Symposium on Discrete Algorithms), ICML 2014 (The 31st International Conference on Machine Learning).

Work Experience

- **Intel Research, Pittsburgh** **May 21, 2007 - August 24, 2007**
  Summer internship at Intel Research, Pittsburgh. **Algorithm Design** to incorporate visual similarity information into the framework of boosted distance metric learning in the Interactive search-assisted diagnosis (ISAD) system. Published at NIPS 2008, SPIE Symposium on Medical Imaging 2007, IJCAI Workshop on Multimodal Information Retrieval 2007, UAI 2007 (Oral presentation), CVPR 2007 (Poster presentation), and CVPR 2008 (Oral presentation).
- **Intel Research, Pittsburgh** **May 22, 2006 - September 1, 2006**
  Summer internship at Intel Research, Pittsburgh. Involved in developing the Interactive Search-Assisted Diagnosis (ISAD) of Medical Images. **Algorithm Design** to identify annotated mammograms from a large medical repository that were similar to the given case. Developed novel algorithms for supervised distance metric learning. Published at AAAI 06 (Oral presentation).
- **Programming and Software Testing** **June, 2002 - August, 2002**
  Summer internship in JinPeng Electronics and Information Device Corporation, Guangzhou Hi-Tech Industrial Development Zone, China.

Courses

- Machine Learning Theory, An Intensive Introduction to Computational Complexity Theory (sit in), Algorithms in the Real World, Advanced Statistical Theory I (audit), Statistical Machine Learning, Machine Learning, Intermediate Statistics, Advanced Probability II: Stochastic Processes, Algorithms, Multimedia Databases and Data Mining, Information Retrieval, Computer Vision, Information Extraction, Statistical Signal Processing, Digital Communication, Neutral

Network, Artificial Intelligence and Pattern Recognition, Advanced Computer Networking and Communications, Technically Speaking.

Awards

- Intel Support for Graduate Study, 2007
- Prize in the Mathematical Modeling Competition held by Huazhong University of Science and Technology in April, 2002.
- Outstanding Graduate of Huazhong University of Science and Technology, 2003.
- Advanced Computer Software Engineer Certificate issued by Ministry of Information Industry, China, 2003.
- Department Distinguished student of Huazhong University of Science and Technology, 2002.

Programming Skills

R, C++, Java, C#, C, Perl, Python, Matlab, and Unix shell scripts.