

Curriculum Vitae

Maria-Florina Balcan

Personal Information

Last name Balcan
First name Maria-Florina
Maiden name Popa
Cell Phone 609 – 712 – 0569
E-mail ninamf@cs.cmu.edu
URL www.cs.cmu.edu/~ninamf/

Research Interests

Computational and Statistical Machine Learning, Algorithmic Aspects in Economics and Game Theory, Algorithms, Theoretical Models for Signal Processing, Artificial Intelligence, Computer Vision.

Current Position

2008 - 2009, Postdoctoral Researcher at Microsoft Research New England, Cambridge, MA.

Education

Ph.D. 2002 - 2008, Carnegie Mellon University, Pittsburgh, PA. Computer Science Department.

Advisor: Avrim Blum.

Dissertation: *New Theoretical Frameworks for Machine Learning*.

Thesis Committee: Avrim Blum, Manuel Blum, Yishay Mansour, Tom Mitchell, Santosh Vempala.

M.S. 2000 – 2002, University of Bucharest, Romania, Faculty of Mathematics and Computer Science, M.S. Degree in Computer Science, GPA 10.00/10.00.

B.S. 1996 – 2000, University of Bucharest, Romania, Faculty of Mathematics, Computer Science Department, major in both Computer Science and Mathematics, GPA 10.00/10.00, “Summa Cum Laude” Diploma.

Honors and Awards

2008, IBM Goldstine Postdoctoral Fellowship in Mathematical Sciences, 2008 (received but declined).

2008, Mark Fulk Best Student Paper Award, 21st Annual Conference on Learning Theory (COLT 2008).

2007 – 2008, IBM Fellowship.

2000 – 2001, Romanian Government Merit Fellowship (during my MS studies).

1996 – 2000, Romanian Government Merit Fellowship (during my undergraduate studies).

October 2001, scholarship offered by World Bank, at Laboratoire d'Analyse et d'Architecture des Systèmes, Dependable Computing and Fault Tolerance (TSF) Team, CNRS, Toulouse, France.

October 1999 - January 2000, European Union Erasmus/Socrates scholarship to study at the University of Patras, Greece.

Other Work Experience

- July 2006 – September 2006, Research Intern, Yahoo! Research, Sunnyvale, CA. Research in Machine Learning and Mechanism Design.
- April 2006, January 2006, and May 2005, Research Visitor, Microsoft Research Silicon Valley, Mountain View, CA. Research in Mechanism Design and Machine Learning.
- August – September 2005, Research Intern, Toyota Technological Institute at Chicago. Research in Machine Learning.
- June – July 2005, Research Intern, IBM T.J. Watson Research Center, Hawthorne, NY. Research in Machine Learning.
- 2000 – 2002, Instructor, Computer Science Department, Faculty of Mathematics, University of Bucharest, Romania.
- October 1999 – January 2000, Research Assistant, Speech and Language Processing Group, Wire Communications Laboratory, University of Patras, Greece. Research in Optical Character Recognition.

Selected Publications

Books and Chapters

- An Augmented PAC Model for Semi-Supervised Learning**, Maria-Florina Balcan and Avrim Blum, Book Chapter in "Semi-Supervised Learning", Chapelle, O., Zien, A., and Scholkopf, B. (Eds.), MIT Press, 2006.
- Search and Knowledge Representation in Artificial Intelligence. Theory and Applications**, Florentina Hristea and Maria-Florina Balcan, University of Bucharest Publishing House, 2005 (in Romanian).

Journal Papers

- Agnostic Active Learning**, Maria-Florina Balcan, Alina Beygelzimer, and John Langford. *Journal of Computer and System Sciences*, 75 (1): 78 – 89, 2009. Special Issue on Learning Theory. **(Invited)**
- The True Sample Complexity of Active Learning**, Maria-Florina Balcan, Steve Hanneke, and Jennifer Wortman. Invited to appear in a special issue of *Machine Learning Journal* for COLT 2008.
- On a Theory of Learning with Similarity Functions**, Maria-Florina Balcan, Avrim Blum, and Nathan Srebro. *Machine Learning Journal*, 72(1 – 2): 89 – 112, 2008. Special issue of *Machine Learning Journal* for COLT 2007. **(Invited)**
- Robust Reductions from Ranking to Classification**, Maria-Florina Balcan, Nikhil Bansal, Alina Beygelzimer, Don Coppersmith, John Langford, and Gregory B. Sorkin. *Machine Learning Journal*, 72 (1 – 2) : 139 – 153, 2008. Special issue of *Machine Learning Journal* for COLT 2007. **(Invited)**
- Reducing Mechanism Design to Algorithm Design via Machine Learning**, Maria-Florina Balcan, Avrim Blum, Jason D. Hartline, and Yishay Mansour, *Journal of Computer and System Sciences*, 74(8): 1245 – 1270, 2009. Special Issue on Learning Theory. **(Invited)**
- Approximation Algorithms and Online Mechanisms for Item Pricing**, Maria-Florina Balcan and Avrim Blum, *Theory of Computing*, 3/9 : 179 – 195, 2007.
- Kernels as Features: On Kernels, Margins, and Low-dimensional Mappings**, Maria-Florina Balcan, Avrim Blum, and Santosh Vempala, *Machine Learning Journal*, 65(1):79 – 94, 2006.
- Approaches to Handwritten/Machine Printed Discrimination Problem**, Maria-Florina Popa and Doru-Cristian Balcan, University of Bucharest Annals, Computer Science, 2000.

Conference Papers

- Finding Low Error Clusterings**, Maria-Florina Balcan and Mark Braverman, Proceedings of the 22nd Annual Conference on Learning Theory (COLT) 2009.
- Price of Uncertainty**, Maria-Florina Balcan, Avrim Blum, and Yishay Mansour, Proceedings of the Tenth ACM Conference on Electronic Commerce (EC) 2009.
- Approximate Clustering without the Approximation**, Maria-Florina Balcan, Avrim Blum, and Anupam Gupta, Proceedings of the ACM-SIAM Symposium on Discrete Algorithms (SODA) 2009.
- Improved Equilibria via Public Service Advertising**, Maria-Florina Balcan, Avrim Blum, and Yishay Mansour, Proceedings of the ACM-SIAM Symposium on Discrete Algorithms (SODA) 2009.
- A Discriminative Framework for Clustering via Similarity Functions**, Maria-Florina Balcan, Avrim Blum, and Santosh Vempala. Proceedings of the 40th ACM Symposium on Theory of Computing (STOC) 2008.
- Improved Guarantees for Learning via Similarity Functions**, Maria-Florina Balcan, Avrim Blum, and Nathan Srebro, Proceedings of the 21th Annual Conference on Learning Theory (COLT) 2008.
- Item Pricing for Revenue Maximization**, Maria-Florina Balcan, Avrim Blum, and Yishay Mansour, Proceedings of the The Ninth ACM Conference on Electronic Commerce (EC) 2008.
- The True Sample Complexity of Active Learning**, Maria-Florina Balcan, Steve Hanneke, and Jennifer Wortman. Proceedings of the 21th Annual Conference on Learning Theory (COLT) 2008.
Received **Mark Fulk Best Student Paper Award**.
- Clustering with Interactive Feedback**, Maria-Florina Balcan and Avrim Blum, Proceedings of the 19th International Conference on Algorithmic Learning Theory (ALT) 2008.
- Margin Based Active Learning**, Maria-Florina Balcan, Andrei Broder, and Tong Zhang, Proceedings of the 20th Annual Conference on Learning Theory (COLT) 2007.
- Robust Reductions from Ranking to Classification**, Maria-Florina Balcan, Nikhil Bansal, Alina Beygelzimer, Don Coppersmith, John Langford, and Gregory B. Sorkin. Proceedings of the 20th Annual Conference on Learning Theory (COLT) 2007.
- A Theory of Loss-leaders: Making Money by Pricing below Cost**, Maria-Florina Balcan, Avrim Blum, Hubert Chan, and MohammadTaghi Hajiaghayi, Proceedings of the 3rd International Workshop on Internet and Network Economics (WINE) 2007.
- Agnostic Active Learning**, Maria-Florina Balcan, Alina Beygelzimer, and John Langford. Proceedings of the 23rd International Conference on Machine Learning (ICML) 2006.
- Approximation Algorithms and Online Mechanisms for Item Pricing**, Maria-Florina Balcan and Avrim Blum. Proceedings of the the Seventh ACM Conference on Electronic Commerce (EC) 2006.
- On a Theory of Learning with Similarity Functions**, Maria-Florina Balcan and Avrim Blum, Proceedings of the 23rd International Conference on Machine Learning (ICML), 2006,
- Mechanism Design via Machine Learning**, Maria-Florina Balcan, Avrim Blum, Jason D. Hartline, and Yishay Mansour. Proceedings of the 46th Annual Symposium on Foundations of Computer Science (FOCS) 2005.
- A PAC-style Model for Learning from Labeled and Unlabeled Data**, Maria-Florina Balcan and Avrim Blum, Proceedings of the 18th Annual Conference on Learning Theory (COLT) 2005.

Kernels as Features: On Kernels, Margins, and Low-dimensional Mappings, Maria-Florina Balcan, Avrim Blum, and Santosh Vempala. Proceedings of the 15th International Conference on Algorithmic Learning Theory (ALT) 2004.

Co-Training and Expansion: Towards Bridging Theory and Practice, Maria-Florina Balcan, Avrim Blum, and Ke Yang, Proceedings of the Eighteenth Annual Conference on Neural Information Processing Systems (NIPS) 2004.

Handwritten Text Localization in Skewed Documents, Ergina Kavallieratou, Doru-Cristian Balcan, Maria-Florina Popa, and Nikos Fakotakis, IEEE International Conference on Image Processing (ICIP) 2001.

Short Surveys and Open Problems

Better Guarantees for Sparsest Cut Clustering, Maria-Florina Balcan, Proceedings of the 22nd Annual Conference on Learning Theory (COLT), 2009.

Item Pricing for Revenue Maximization, Maria-Florina Balcan, Avrim Blum, and Yishay Mansour. SIGecom Exchanges, Volume 7.3, 2008. (**Invited**)

Mechanism Design, Machine Learning, and Pricing Problems, Maria-Florina Balcan and Avrim Blum, SIGecom Exchanges, Volume 7.1, special issue on Combinatorial Auctions, 2007. (**Invited**)

Open Problems in Efficient Semi-Supervised PAC Learning, Maria-Florina Balcan and Avrim Blum, Proceedings of 20th Annual Conference on Learning Theory (COLT), 2007.

Selected Workshop Papers

Similarity-Based Theoretical Foundation for Sparse Parzen Window Prediction, Maria-Florina Balcan, Avrim Blum, and Nathan Srebro, ICML/UAI/COLT 2008 Workshop on Sparse Optimization and Variable Selection.

Asymptotic Active Learning, Maria-Florina Balcan, Eyal-Even Dar, Steve Hanneke, Michael Kearns, Yishay Mansour, and Jennifer Wortman, NIPS 2007 Workshop on the Principles of Learning Problem Design.

Person Identification in Webcam Images: An Application of Semi-Supervised Learning, Maria-Florina Balcan, Avrim Blum, Pakyan Choi, John Lafferty, Brian Pantano, Mugizi Robert Rwebangira and Xiaojin Zhu, ICML 2005 Workshop on Learning with Partially Classified Training Data.

Technical Reports

Random Sampling Auctions for Limited Supply, Maria-Florina Balcan, Nikhil Devanur, Jason D. Hartline, and Kunal Talwar. Technical Report, CMU-CS-07-154, 2007.

Manuscripts

Agnostic Clustering, Maria-Florina Balcan, Heiko Röglin and Shang-Hua Teng, 2009.

Other Technical Writings

New Theoretical Frameworks for Machine Learning, Maria-Florina Balcan, Thesis Proposal, Computer Science Department, Carnegie Mellon University, May 2007.

Issues of Search and Knowledge Representation in Artificial Intelligence, Florentina Hristea and Maria-Florina Balcan, Textbook, University of Bucharest Publishing House, 2004 (in Romanian).

Data Defining Language, Doru-Cristian Balcan, Maria-Florina Balcan, and Dorin Paun, chapter in “Procedural and Non-procedural Query Resolution in Oracle8”, University of Bucharest Publishing House, 2002 (in Romanian).

Dependability Modelling and Analysis using Generalized Stochastic Petri Nets, Maria-Florina Popa, MS Thesis, University of Bucharest, July 2002 (in Romanian).

Algorithms and Mathematical Models for Image Processing, Maria-Florina Popa, BS Thesis, University of Bucharest, July 2000 (in Romanian).

New Methods in Handwritten/Machine Printed Discrimination, Maria-Florina Popa and Doru-Cristian Balcan, The Eighth Conference of Applied and Industrial Mathematics (CAIM), 2000.

Selected Other Workshop Papers

Learning with Multiple Similarity Function, Maria-Florina Balcan, Avrim Blum, and Nathan Srebro, NIPS 2008 on Kernel Learning: Automatic Selection of Optimal Kernels.

On a Theory of Kernels as Similarity Functions, Maria-Florina Balcan and Avrim Blum, Mathematical Foundations of Learning Theory-II, 2006.

Sponsored Search Auctions Design via Machine Learning, Maria-Florina Balcan, Avrim Blum, Jason D. Hartline, and Yishay Mansour, ACM-EC 2005 Workshop on Sponsored Search Auctions.

A Sample-Complexity Analysis of Learning from Labeled and Unlabeled Data, Maria-Florina Balcan and Avrim Blum, NIPS 2004 Workshop on (Ab)Use of Bounds.

A PAC-style Model for Learning from Labeled and Unlabeled Data, Maria-Florina Balcan and Avrim Blum, Mathematical Foundations of Learning Theory-I, 2004.

Teaching Experience

Guest Lecturer, Fall 2007, Carnegie Mellon University, for 15 – 892 **Foundations of Electronic Marketplaces** (Professor Tuomas Sandholm).

Instructor, Spring 2007, Carnegie Mellon University, for 15 – 859(B) **Machine Learning Theory**.

Guest Lecturer, Spring 2006, Carnegie Mellon University, for 15 – 859(B) **Machine Learning Theory** (Professor Avrim Blum).

Teaching Assistant, Fall 2004 and Spring 2004, Carnegie Mellon University, for 15 – 451 **Algorithms**.

Teaching Assistant, Spring 2002 and Spring 2001, University of Bucharest, for **Artificial Intelligence**.

Teaching Assistant, Fall 2001, Spring 2001, and Fall 2000, University of Bucharest, for **Introduction to Computer Science**.

Professional Service

Synergistic Activities

Co-Organizer, NIPS 2008 Workshop on New Challenges in Theoretical Machine Learning: Learning with Data-dependent Concept Spaces.

Programm Committee Member

50th Annual Symposium on Foundations of Computer Science (**FOCS**) 2009.

22nd Annual Conference on Learning Theory (**COLT**) 2009.

International Conference on Machine Learning (**ICML**) 2008.

International Conference on Machine Learning (**ICML**) 2007.

Journal Refereeing

Journal of Machine Learning Research, Machine Learning Journal, Theory of Computing, Transactions on Algorithms, IEEE Transactions on Information Theory, Journal of Artificial Intelligence Research, Artificial Intelligence Journal, Neural Computation.

Conference Refereeing

IEEE Foundations of Computer Science (FOCS), ACM Symposium on Theory of Computation (STOC), ACM-SIAM Symposium on Discrete Algorithms (SODA), Conference on Learning Theory (COLT), Neural Information Processing Systems (NIPS), Symposium on Theoretical Aspects of Computer Science (STACS), Annual European Symposium on Algorithms (ESA), International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX), Symposium on Parallelism in Algorithms and Architectures (SPAA), ACM Conference on Electronic Commerce (EC), International Workshop on Internet and Network Economics (WINE), Workshop on Approximation and Online Algorithms (WAOA).

Other Activities

STOC 2008 student travel scholarship awardee.

ACM EC 2007 student travel scholarship awardee.

ICML 2006 and ICML 2003 student volunteer and student travel scholarship awardee.

NIPS 2005 and NIPS 2004 student volunteer.

Selected Invited Talks

The Dynamics of Equilibria

Harvard University, Economics and Computer Science Seminar. May 2009.

A Computational Theory of Clustering

Information Theory and Applications Workshop. February 2009.

A Theory of Learning and Clustering via Similarity Functions

Google Research New York. April 2008.

Ecole Polytechnique Fédérale de Lausanne, School Seminar. April 2008.

Georgia Institute of Technology. April 2008.

University of Washington, CSE Colloquium. April 2008.

Massachusetts Institute of Technology, CS Special Seminar Series. March 2008.

University of Michigan, Ann Arbor, CSE Colloquium. March 2008.

Microsoft Research Silicon Valley. March 2008.

University of Stanford, Computer Science Seminar. March 2008.

University of Southern California, CS Colloquium Series. March 2008.

University of Pennsylvania, Wharton School Statistics Department Seminar. February 2008.

Georgia Institute of Technology, ARC Colloquium. February 2008.

Microsoft Research Redmond. February 2008.

Cornell Theory Seminar. November 2007.

China Theory Week, Tsinghua University. September 2007.

University of Pennsylvania, Machine Learning Lunch. September 2007.

Mechanism Design, Machine Learning, and Pricing Problems

University of Southern California, CS Colloquium Series. December 2007.

Duke Computer Science Colloquia. December 2007.

Brown CS Seminar. November 2007.

INFORMS. November 2006.

University of California, San Diego, Theory Seminar. October 2006.

Yahoo! Research, Sunnyvale, CA. October 2006.

Second Bertinoro Workshop on Algorithmic Game Theory (AGATE 2006). July 2006.

CMU Workshop on Electronic Market Places. May 2006.

Item Pricing for Revenue Maximization in Combinatorial Auctions

Dagstuhl Workshop on Computational Social Systems and the Internet. July 2007.

An Augmented PAC Model for Semi-Supervised Learning

Toyota Technological Institute at Chicago (TTI-C). August 2005.

IBM Research T.J. Watson. June 2005.

Microsoft Research Silicon Valley. May 2005.

Kernels as Features: On Kernels, Margins, and Low-dimensional Mappings

IBM Research T.J. Watson. July 2005.

References

Prof. Avrim Blum

Department of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213 – 3891
E-mail: avrim@cs.cmu.edu

Prof. Yishay Mansour

School of Computer Science
Tel Aviv University
Tel Aviv, ISRAEL
E-mail: mansour.yishay@gmail.com
E-mail: mansour@tau.ac.il

Prof. Santosh Vempala

College of Computing
Georgia Institute of Technology
Atlanta GA 30332 – 0280.
E-mail: vempala@cc.gatech.edu

Prof. Michael Kearns

Department of Computer and Information Science
University of Pennsylvania
Philadelphia, PA 19104 – 6389
E-mail: mkearns@cis.upenn.edu

Prof. Tom Mitchell

Department of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213 – 3891
E-mail: Tom.Mitchell@cmu.edu

Prof. Tong Zhang

Department of Statistics
Rutgers University
Piscataway, NJ 08854 – 8019
E-mail: tzhang@stat.rutgers.edu