

# Satyen Kale

satyen@satyenkale.com  
www.satyenkale.com

## Curriculum Vitae

---

### RESEARCH INTERESTS

---

Design of efficient and practical algorithms for fundamental problems in **machine learning** and **optimization**, specifically:

- decision-making under uncertainty
- combinatorial optimization
- statistical learning theory
- convex optimization

---

### EDUCATION

---

**Princeton University, M.A. (2004) and Ph.D. (2007), Computer Science**

Thesis: Efficient Algorithms Using the Multiplicative Weights Updates Method. Advised by Prof. Sanjeev Arora.

**Indian Institute of Technology Bombay, B.Tech. (2002), Computer Science and Engineering**

Thesis: Spectral Algorithms for Data Representation and Manipulation. Advised by Prof. Abhiram Ranade. Graduated 2<sup>nd</sup> in class with a GPA of 9.85/10.00.

**Indian Institute of Statistics, Calcutta, summers 1999-2002**

Participant in Nurture Program in Mathematics, equivalent of a Master's course in mathematics.

---

### WORK EXPERIENCE

---

**Google Research, New York, NY, Senior Research Scientist, 2016-present**

Working on large scale machine learning research and projects.

**Yahoo! Research, New York, NY, Senior Research Scientist, 2013-2016**

Worked on large scale machine learning projects related to spam filtering and click through rate estimation. Conducting research on scalable machine learning.

**Columbia University, New York, NY, Adjunct Professor, Spring 2016**

Teaching an introductory course on Machine Learning.

**IBM Research, Yorktown Heights, NY, Research Staff Member, 2011-2013**

Worked on data analysis projects and conducted research on machine learning and optimization.

**Yahoo! Research, Santa Clara, CA, Postdoctoral Researcher, 2009-2011**

Designed and implemented algorithms for display advertising inventory management, supply forecasting, and spam classification. Conducted research on machine learning and game theory.

**Microsoft Research Redmond and New England, Postdoctoral Researcher, 2007-2009**

Conducted research on machine learning, sublinear time algorithms, and complexity theory. Designed and implemented advertisement selection algorithms for Bing search queries.

**Microsoft Research Silicon Valley Center, Mountain View, CA, Summer Intern, 2006**

Researched privacy-preserving algorithms for releasing contingency tables, learning, and online optimization.

**IBM Almaden Research Center, San Jose, CA, Summer Intern, 2005**

Researched streaming algorithms, game theory, and metric embeddings.

**ETH Zürich, Switzerland, Summer Intern, 2001**

Designed and implemented sieve methods to find smooth numbers in short intervals.

---

## PUBLICATIONS

---

PEER-REVIEWED CONFERENCE AND JOURNAL PAPERS

- [1] **Online Sparse Linear Regression**  
D. Foster, S. Kale, and H. Karloff. *Conference on Learning Theory (COLT)*, 2016. Gave talk at conference.
- [2] **Online Gradient Boosting**  
A. Beygelzimer, E. Hazan, S. Kale, and H. Luo. *Neural Information Processing Systems (NIPS)*, 2015.
- [3] **Optimal and Adaptive Algorithms for Online Boosting**  
A. Beygelzimer, S. Kale, and H. Luo. *International Conference on Machine Learning (ICML)*, 2015. **Best paper award.**
- [4] **Budgeted Prediction With Expert Advice**  
K. Amin, S. Kale, G. Tesauro, and D. Turaga. *Conference of Association for the Advancement of Artificial Intelligence (AAAI)*, 2015.
- [5] **Beyond the Regret Minimization Barrier: an Optimal Algorithm for Stochastic Strongly-Convex Optimization**  
E. Hazan and S. Kale. *Journal of Machine Learning Research (JMLR)*, 2014. Preliminary version in *Conference on Learning Theory (COLT)*, 2011. Gave talk at conference.
- [6] **Taming the Monster: A Fast and Simple Algorithm for Contextual Bandits**  
A. Agarwal, D. Hsu, J. Langford, S. Kale, L. Li, and R. E. Schapire. *International Conference on Machine Learning (ICML)*, 2014.
- [7] **Multiarmed Bandits With Limited Expert Advice**  
S. Kale. *Conference on Learning Theory (COLT)*, 2014. Gave talk at conference.
- [8] **Adaptive Market Making via Online Learning**  
J. Abernethy and S. Kale. *Neural Information Processing Systems (NIPS)*, 2013.
- [9] **Bargaining for Revenue Shares on Tree Trading Networks**  
A. Ghosh, S. Kale, K. Lang and B. Moseley. *International Joint Conference on Artificial Intelligence (ICJAI)*, 2013.
- [10] **The Approximability of the Binary Paintshop Problem**  
A. Gupta, S. Kale, V. Nagarajan, R. Saket and B. Schieber. *International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX)*, 2013.
- [11] **Noise Tolerance of Expanders and Sublinear Expander Reconstruction**  
S. Kale, Y. Peres, and C. Seshadhri. *SIAM Journal on Computing (SICOMP)*, 2013. Preliminary version in *Foundations of Computer Science (FOCS)*, 2008.

- [12] **Near-Optimal Algorithms for Online Matrix Prediction**  
E. Hazan, S. Kale and S. Shalev-Shwartz. *Conference on Learning Theory (COLT)*, 2012. Gave talk at conference.
- [13] **Projection-free Online Learning**  
E. Hazan and S. Kale. *International Conference on Machine Learning (ICML)*, 2012.
- [14] **Efficient and Practical Stochastic Subgradient Descent for Nuclear Norm Regularization**  
H. Avron, S. Kale, S. P. Kasiviswanathan, and V. Sindhvani. *International Conference on Machine Learning (ICML)*, 2012.
- [15] **Contextual Bandit Learning with Predictable Rewards**  
A. Agarwal, M. Dudik, S. Kale, J. Langford and R. E. Schapire. *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2012.
- [16] **The Multiplicative Weights Update method: a Meta-Algorithm and some Applications**  
S. Arora, E. Hazan, and S. Kale. *Theory of Computing Journal*, 2012.
- [17] **An Online Portfolio Selection Algorithm with Regret Logarithmic in Price Variation**  
E. Hazan and S. Kale. *Mathematical Finance*, 2012. Preliminary version appeared as **On Stochastic and Worst-case Models for Investing** in *Neural Information Processing Systems (NIPS)*, 2009.
- [18] **Newtron: an Efficient Bandit algorithm for Online Multiclass Prediction**  
E. Hazan and S. Kale. *Neural Information Processing Systems (NIPS)*, 2011.
- [19] **Efficient Optimal Learning for Contextual Bandits**  
M. Dudik, D. Hsu, S. Kale, N. Karampatziakis, J. Langford, L. Reyzin, and T. Zhang. *Conference on Uncertainty in Artificial Intelligence (UAI)*, 2011.
- [20] **Who Moderates the Moderators?: Crowdsourcing Abuse Detection in User-Generated Content**  
A. Ghosh, S. Kale, and R. P. McAfee. *Conference on Electronic Commerce (EC)*, 2011.
- [21] **Cross-Validation and Mean-Square Stability**  
S. Kale, R. Kumar and S. Vassilvitskii. *Innovations in Computer Science (ICS)*, 2011.
- [22] **Combinatorial Approximation Algorithms for MaxCut Using Random Walks**  
S. Kale and C. Seshadhri. *Innovations in Computer Science (ICS)*, 2011.
- [23] **An Expansion Tester for Bounded Degree Graphs**  
S. Kale and C. Seshadhri. *SIAM Journal on Computing (SICOMP)*, 2011. Preliminary version in *International Colloquium on Automata, Languages and Programming (ICALP)*, 2008.
- [24] **Better Algorithms for Benign Bandits**  
E. Hazan and S. Kale. *Journal of Machine Learning Research (JMLR)*, 2011. Preliminary version in *Symposium on Discrete Algorithms (SODA)*, 2009.
- [25] **Non-Stochastic Bandit Slate Problems**  
S. Kale, L. Reyzin and R. E. Schapire. *Neural Information Processing Systems (NIPS)*, 2010.
- [26] **Learning Rotations with Little Regret**  
E. Hazan, S. Kale and M. K. Warmuth. *Conference on Learning Theory (COLT)*, 2010. Gave talk at conference.

- [27] **Extracting Certainty from Uncertainty: Regret Bounded by Variation in Costs**  
E. Hazan and S. Kale. *Machine Learning Journal* (MLJ), 2010, special issue for COLT 2008. Preliminary version in *Conference on Learning Theory* (COLT), 2008.
- [28]  **$O(\sqrt{\log n})$  Approximation to SPARSEST CUT in  $\tilde{O}(n^2)$  Time**  
S. Arora, E. Hazan, and S. Kale. Accepted to *SIAM Journal on Computing* (SICOMP), 2010. Preliminary version in *Foundations of Computer Science* (FOCS), 2004. Gave talk at conference.
- [29] **Online Submodular Minimization**  
E. Hazan and S. Kale. *Neural Information Processing Systems* (NIPS), 2009.
- [30] **The Uniform Hardcore Lemma via Approximate Bregman Projections**  
B. Barak, M. Hardt, and S. Kale. *Symposium on Discrete Algorithms* (SODA), 2009.
- [31] **Computational Equivalence of Fixed Points and No Regret Algorithms, and Convergence to Equilibria**  
E. Hazan and S. Kale. *Conference on Neural Information Processing Systems* (NIPS), 2007.
- [32] **A Combinatorial, Primal-Dual approach to Semidefinite Programs**  
S. Arora and S. Kale. *Symposium of Theory of Computing* (STOC), 2007. Gave talk at conference.
- [33] **Privacy, Accuracy, and Consistency Too: A Holistic Solution to Contingency Table Release**  
B. Barak, K. Chaudhuri, C. Dwork, S. Kale, F. McSherry, and K. Talwar. *Symposium on Principles of Database Systems* (PODS), 2007.
- [34] **Efficient Aggregation Algorithms for Probabilistic Data**  
T. S. Jayram, S. Kale, and E. Vee. *Symposium on Discrete Algorithms* (SODA), 2007.
- [35] **A Variation on SVD Based Image Compression**  
A. Ranade, S. S. Mahabalarao, and S. Kale. *Image and Vision Computing*, 2007. Preliminary version in *Workshop on Computer Vision, Graphics, and Image Processing* (WCVGIP), 2006.
- [36] **Logarithmic Regret Algorithms for Online Convex Optimization**  
E. Hazan, A. Agarwal, and S. Kale. *Machine Learning Journal* (MLJ), 2007, special issue for COLT 2006. Preliminary version with A. Kalai in *Conference on Learning Theory* (COLT), 2006.
- [37] **Algorithms for Portfolio Management Based on the Newton Method**  
A. Agarwal, E. Hazan, S. Kale, and R. E. Schapire. *International Conference on Machine Learning* (ICML), 2006.
- [38] **A Fast Random Sampling Algorithm for Sparsifying Matrices**  
S. Arora, E. Hazan, and S. Kale. *Workshop on Randomization and Computation* (RANDOM), 2006.
- [39] **Fast Algorithms for Approximate Semidefinite Programming Using the Multiplicative Weights Update method**  
S. Arora, E. Hazan, and S. Kale. *Foundations of Computer Science* (FOCS), 2005. Gave talk at conference.
- [40] **Analysis and Algorithms for Content-based Event Matching**  
S. Kale, E. Hazan, F. Cao and J. P. Singh. *Workshop on Distributed Event-Based Systems* (DEBS), 2005.

## TECHNICAL REPORTS

- [1] **Efficient Algorithms Using the Multiplicative Weights Update Method**  
S. Kale. Ph.D. Thesis. Princeton Technical Report TR-804-07, 2007.
- [2] **Approximating Quadratic Programs with Positive Semidefiniteness Constraints**  
E. Hazan and S. Kale. Princeton Technical Report TR-746-06, 2004.

---

## INVITED TALKS

---

### **Online Boosting Algorithms**

- National University of Singapore, Singapore, November 2015.
- Courant Institute of Mathematical Sciences, New York University, New York, NY, October 2015.
- Department of Computer Science, Cornell University, Ithaca, NY, September 2015.

### **Efficient Optimal Learning for Contextual Bandits**

- Departments of Computer and Information Sciences, University of Pennsylvania, Philadelphia, PA, February 2015.
- Lorentz Center, Leiden University, Netherlands, November 2014.
- Microsoft Research, Redmond, WA, May 2014.
- Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, September 2013.
- Janelia Conference, Ashburn, VA, May 2012.
- Machine Learning Workshop, University of Minnesota, Minneapolis, MN, March 2012.

### **Who Moderates the Moderators? Crowdsourcing Abuse Detection in User-Generated Content**

- Workshop on Workshop on Transactional Machine Learning and E-Commerce, NIPS 2014.

### **The Matrix Multiplicative Weights Algorithm: Applications to SDP and Online Matrix Prediction**

- ICERM, Brown University, Providence, RI, February 2014.

### **Efficient First-Order Methods for Large-Scale Stochastic Convex Optimization**

- ExxonMobil Research, Annandale, NJ, April 2014.
- Algorithmic Frontiers Workshop, EPFL, Lausanne, Switzerland, June 2012.

### **Near-Optimal Algorithms for Online Matrix Prediction**

- Workshop on Large Scale Matrix Analysis and Inference, NIPS 2013.

### **A Combinatorial, Primal-Dual Approach to Semidefinite Programs**

- Approximation Algorithms workshop, Princeton, NJ, June 2011.
- HPOPT 2010, Tilburg University, The Netherlands, June 2010.
- MMDS 2008, Stanford, CA, June 2008.
- IBM Almaden Research Center, San Jose, CA, February 2008.
- INFORMS 2007, Seattle, WA, November 2007.
- Computer Science department, University of Washington, Seattle, WA, October 2007.

### **Combinatorial Approximation Algorithms for MaxCut Using Random Walks**

- IPAM workshop on Discrete Optimization, Los Angeles, CA, October 2010

### **Noise Tolerance of Expanders and Sublinear Expander Reconstruction**

- Computer Science department, University of Washington, Seattle, WA, November 2009.
- Computer Science department, Princeton University, Princeton, NJ, November 2008
- MIT Applied Math Seminar, Cambridge, MA, October 2008.
- Microsoft Research, Redmond, WA, May 2008.

### **Fast Algorithms for Approximate Semidefinite Programming Using the Multiplicative Weights Update Method**

- ISMP 2006, Rio de Janeiro, Brazil, August 2006.
- INFORMS 2005, San Francisco, CA, November 2005.

### **$O(\sqrt{\log n})$ Approximation to SPARSEST CUT in $\tilde{O}(n^2)$ Time**

- INFORMS 2008, Washington, D.C., October 2008.
- DIMACS Mixer Series, Rutgers University, Piscataway, NJ, September 2004.

### **The Multiplicative Weights Update Method**

- IBM T.J. Watson Research Center, Yorktown Heights, NY, May 2006.
- Yahoo! Research, Sunnyvale, CA, July 2005.
- Computer Science department, Princeton University, Princeton, NJ, October 2004.

---

## PATENT APPLICATIONS

---

### **Methods and Systems for Ad Placement Planning**

E. Vee, D. Chen, P. Chen, S. Kale, S. Mandalapu, and C. Nagarajan. U.S. Patent US9135632 B2.

### **Predicting Web Advertisement Click Success By Using Head-To-Head Ratings**

M. Bayati, M. Braverman, S. Kale, and Y. Makarychev. U.S. Patent Application 20100198685.

### **Consistent Contingency Table Release**

B. Barak, K. Chaudhuri, C. Dwork, S. Kale, F. McSherry, and K. Talwar. U.S. Patent Application 20090182797.

---

## TEACHING EXPERIENCE

---

### **Lecturer for COMS 4771: *Introduction to Machine Learning*, Spring 2016**

Computer Science Department, Columbia University, New York, NY.

### **Guest lecturer for COS 521: *Advanced Algorithm Design*, Fall 2005, 2006**

Computer Science Department, Princeton University, Princeton, NJ.

### **Teaching Assistant for COS 226: *Algorithms and Data Structures*, Fall 2004**

Computer Science Department, Princeton University, Princeton, NJ.

### **Teaching Assistant for COS 341: *Discrete Mathematics*, Fall 2003**

Computer Science Department, Princeton University, Princeton, NJ.

---

## HONORS

---

**Best paper award at International Conference on Machine Learning, 2015**

“Optimal and Adaptive Algorithms for Online Boosting” awarded best paper at ICML 2015.

**Bronze medal in 37<sup>th</sup> International Mathematics Olympiad, 1997**

Represented India at the IMO in Mar Del Plata, Argentina.

**Scholarship in National Talent Search Examination, 1996**

One of 750 students are awarded scholarships every year all over India.

---

## PROFESSIONAL SERVICE

---

**Ph.D. committee member:** Haipeng Luo (Princeton University, 2008), Qiang Ma (Rutgers University, 2008).

**Program committee member:** International Conference on Machine Learning (ICML) 2009, 2012 (area chair), 2013 (area chair), and 2016 (area chair) Conference on Learning Theory (COLT) 2011, 2012, 2013, 2014, 2015, 2016 and 2017, Symposium on Discrete Algorithms (SODA) 2011, and Conference on Knowledge Discovery and Mining (SIGKDD) 2010.

**Conference organization:** Serving as COLT 2017 program chair. Served as COLT 2016 local chair and COLT 2015 publications chair. Organized sessions on *Online Learning* at INFORMS 2009 and *Semidefinite Programming and Applications* at INFORMS 2008.

**Papers referee for journals and conferences:** Journal of the ACM (JACM), SIAM Journal on Computing (SICOMP), Journal of Machine Learning Research (JMLR), Machine Learning Journal (MLJ), Algorithmica, Theoretical Computer Science journal (TCS), Transactions on Algorithms (TALG), and Statistics and Computing journal (STCO), Foundations of Computer Science (FOCS), Symposium on Theory of Computation (STOC), Symposium on Discrete Algorithms (SODA), Conference on Learning Theory (COLT), Innovations in Theoretical Computer Science (ITCS), International Conference on Machine Learning (ICML), Neural Information Processing Systems (NIPS), ACM Conference on Knowledge Discovery and Mining (SIGKDD), Conference on Uncertainty in Artificial Intelligence (UAI), Conference on Integer Programming and Combinatorial Optimization (IPCO), International Conference on Very Large Data Bases (VLDB), Symposium on Theoretical Aspects of Computer Science (STACS), International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX), International Workshop on Randomization and Computation (RANDOM), and International Symposium on Information Theory (ISIT).

---

## PROGRAMMING LANGUAGES

---

Scala, Java, C, C++, Matlab, Python