

Nakul Verma

CONTACT INFORMATION

Columbia University
Computer Science Dept. (MC 0401)
500 W120 St, New York, NY 10027

email: verma@cs.columbia.edu
www: www.cs.columbia.edu/~verma
phone: (212) 939-7103

EDUCATION

University of California, San Diego, La Jolla, California, USA

Doctor of Philosophy, Computer Science, September 2012.

- Advisor: Professor Sanjoy Dasgupta
- Area of Study: Machine Learning

Master of Science, Computer Science, June 2008.

Bachelor of Science, Computer Science, December 2004.

AWARDS

Research & professional service related

- ICML reviewer award, 2015.
- Best paper, Wireless Health 2012.

Teaching related

- Janelia Teaching Fellowship, 2015.

Programming related

- Second place in UCSD Programming Contest, 2005.
- Fourth place in Southern California Regional ACM Programming Contest, 2004.
- Third place in UCSD Programming Contest, 2004. (top undergraduate finisher)
- Fourth place in UCSD Datamining Contest, 2004.

University of California, San Diego

- Provost honors, 2001–2004.

TEACHING EXPERIENCE

Columbia University, New York City, New York USA

Lecturer in Discipline

Fall 2017 – Present

- Machine Learning (COMS 4771) – Fall: 17, 18, Spring:18, 19, Summer:15, 18.
- Methods in Unsupervised Learning (COMS 4995) – Fall: 18, Summer: 18
- Automata and Complexity Theory (COMS 3261) – Fall: 17

Adjunct Assistant Professor

Summer 2015

- Taught Machine Learning course to graduate and undergraduate students.
- Responsible to designing the lectures, homeworks and exams.

University of California, San Diego, La Jolla, California USA

Teaching Assistant and Course Tutor

2003–2007

- Advanced Data Structures – WI04, SP05, SP06, FA06
- Compiler Construction – WI05, FA05, WI06, FA06, WI07
- Design and Analysis of Algorithms – SP05
- Discrete Mathematics – SP04
- Intro to Computer Science and OOP: Java – WI03, SP03, FA03, FA04, WI07
- California State Summer School for Math and Science (COSMOS) program, Computer science cluster – SU06

Columbia University, New York City, New York USA

Lecturer in Discipline

Fall 2017 – Present

- Heading research work in understanding and exploiting the intrinsic structure in data (eg. manifold or sparse structure) to design effective machine learning algorithms.

Janelia Research Campus, Ashburn, Virginia USA

Research Specialist

2013–2017

- Developing biologically plausible statistical models to understand organism behavior from observational data.
- Conduct cutting-edge machine learning research to understand the mathematical structure of behavior.

Amazon.com Inc., Seattle, Washington USA

Research Scientist

2012–2013

- Developing and analyzing automated models for detecting fraudulent transactions in real time.
- Part of the fraud analytics team; responsible for all live European digital orders.

Yahoo! Labs, Bangalore, India

Research Intern

Summer 2011

- Developed learning models for improved classification and visualization for hierarchical multi-class data.

Qualcomm Incorporated, San Diego, California USA

Research Intern

Summer 2008

- Applying machine learning and data mining concepts to design provably correct algorithms for phone personalization in a memory constraint environment.
- Part of the Advanced Technologies Group in the Office of the Chief Scientist.

University of California, San Diego, La Jolla, California USA

Graduate Student Researcher

2010–2012

- Part of an interdisciplinary project, called CitiSense, whose goal is to monitor air pollution via mobile sensors and study its effects on the local community.
- Developing spatio-temporal models to learn and predict pollution levels in the San Diego county.

Graduate Student Researcher

2005–2012

- Worked with Prof. Sanjoy Dasgupta on learning and prediction problems on data that has low-intrinsic dimension.

Undergraduate Student Researcher

2003

- Part of Prof. Alex Orailoglu's Reliable System Synthesis Group, which focuses on testing and verification of VLSI circuits.
- Assisted a PhD student in reducing power consumption issues during the chip-testing phase.

RESEARCH
GRANTS

Alliance Joint Project Grant

co-Principal Investigator

2019-20

- Investigating data-driven methodologies to automatically discover physical laws in mechanics and astrophysics data.
- Received seed-funding to conduct high-impact research and foster transatlantic collaborations with researchers.
- Funding amount: \$15,000.

Data Science Institute (DSI) Scholars Program

Faculty Mentor

Spring 2019

- Investigating non-traditional embedding techniques for modern large-scale data.
- Awarded funding to support students to work on two separate research projects.
- Funding amount: \$5,000.

PROFESSIONAL
ACTIVITIES

Reviewer and/or program committee member:

- Nature Methods
- Journal of Machine Learning Research (JMLR)
- Machine Learning Journal (MLJ)
- IEEE Transactions on Information Theory (T-IT)
- Discrete and Computational Geometry (DCG)
- Conference on Learning Theory (COLT)
- Neural Information Processing Systems (NIPS)
- International Conference on Machine Learning (ICML)
- International Conference on Artificial Intelligence and Statistics (AISTATS)
- Association for the Advancement of Artificial Intelligence (AAAI)
- Neural Computation
- IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)
- IEEE Transactions on Signal Processing (TSP)
- IEEE Signal Processing Letters (SPL)
- International Journal of Computer Vision (IJCV)
- EURASIP Journal of Wireless Communications and Networking (JWCN)
- ACM SIGKDD Conference on Knowledge Data Discovery and Data Mining (KDD) Learning from Small Sample Sizes.
- IEEE International Conference on Data Mining (ICDM) High Dimensional Mining.
- International Conference on Distributed Computing and Networking (ICDCN) Big Data Analytics.

PUBLICATIONS

Discovering the vocabulary of mouse social behavior via metric learning. *(In Progress)*

High-throughput online prediction of zebrafish seizure activity. *(In Progress)*

M. Aalto, N. Verma. Metric Learning on Manifolds. *Under review: International Conference on Machine Learning (ICML), preprint: arXiv:1902.01738, 2019.*

A. Lamy, Z. Zhong, A. Menon, N. Verma. Noise-tolerant fair classification. *Under review: International Conference on Machine Learning (ICML), preprint: arXiv:1901.10837, 2019.*

J. Im, N. Verma, K. Branson. Stochastic Neighbor Embeddings under f -divergences. *Under review: International Conference on Machine Learning (ICML), preprint: arXiv:1811.01247, 2018.*

- B. Sauerbrei, J. Guo, J. Zheng, W. Guo, M. Kabra, N. Verma, K. Branson, A. Hantman. The cortical dynamics orchestrating skilled prehension. *Under review: Cell, preprint: biorXiv doi:10.1101/266320*, 2018.
- S. Kpotufe, N. Verma. Time-accuracy tradeoffs in Kernel prediction: controlling prediction quality. *Journal of Machine Learning Research (JMLR)*, 2017.
- N. Verma, K. Branson. Sample complexity of learning Mahalanobis distance metrics. *Neural Information Processing Systems (NIPS)*, 2015.
- N. Verma. Distance preserving embeddings for general n -dimensional manifolds. *Journal of Machine Learning Research (JMLR)*, 2013.
- B. Milosevic, J. Yang, N. Verma, S. Tilak, P. Zappi, E. Farella, L. Benini, T. Rosing. Efficient energy management and data recovery in sensor networks using latent variables based tensor factorization. *Sixteenth IEEE International Conference on Modelling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*, 2013.
- K. Patrick, B. Griswold, G. Norman, F. Raab, J. Kerr, B. Demchak, I. Krueger, S. Dasgupta, N. Nikzad, C. Ziftci, N. Verma, P. Zappi, E. Bales, N. Quick. Mobile Technologies and the Exposome: Continuous Assessment of Environmental Exposures Critical to Health. *Lecture Notes on Wireless Healthcare Research*, pp. 25–32, 2013.
- N. Verma. Learning from data with low intrinsic dimension. Ph.D. thesis, *Dept. of Computer Science and Engineering, UC San Diego*, 2012.
- N. Verma. Distance preserving embeddings for low-dimensional manifolds. *Conference on Learning Theory (COLT)*, 2012.
- N. Verma, D. Mahajan, S. Sellamanickam, V. Nair. Learning hierarchical similarity metrics. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
- N. Nikzad, N. Verma, C. Ziftci, E. Bales, N. Quick, P. Zappi, K. Patrick, S. Dasgupta, I. Krueger, T. Rosing, W. Griswold. CitiSense: Improving geospatial environmental assessment of air quality using a wireless personal exposure monitoring system. *Wireless Health*, 2012. (**best paper**)
- C. Ziftci, N. Nikzad, N. Verma, P. Zappi, E. Bales, I. Krueger, W. Griswold. Citisense: mobile air quality sensing for individuals and communities. *Conference on Systems, Programming, and Applications: Software for Humanity*, 2012.
- N. Verma. A note on random projections for preserving paths on a manifold. *UC San Diego Tech. Report CS2011-0971*, 2011.
- N. Verma, P. Zappi, T. Rosing. Latent variables based data estimation for sensing applications. *IEEE International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, 2011.
- B. Babenko, N. Verma, P. Dollar, S. Belongie. Multiple instance learning with manifold bags. *Twenty-Eighth International Conference on Machine Learning (ICML)*, 2011.
- N. Nikzad, C. Ziftci, P. Zappi, N. Quick, P. Aghera, N. Verma, B. Demchak, K. Patrick, H. Shacham, T. Rosing, I. Krueger, W. Griswold, S. Dasgupta. CitiSense - Adaptive

Services for Community-Driven Behavioral and Environmental Monitoring to Induce Change. *UC San Diego Tech. Report CS2011-0961*, 2011.

N. Verma, S. Kpotufe, S. Dasgupta. Which spatial trees are adaptive to intrinsic dimension?. *Twenty-Fifth Conference on Uncertainty in Artificial Intelligence (UAI)*, 2009.

N. Verma. Mathematical advances in manifold learning. *UC San Diego Tech. Report*, 2008.

Y. Freund, S. Dasgupta, M. Kabra, N. Verma. Learning the structure of manifolds using random projections. *Twenty-First Conference on Neural Information Processing Systems (NIPS)*, 2007.

S. Dasgupta, D. Hsu, N. Verma. A concentration theorem for projections. *Twenty-Second Conference on Uncertainty in Artificial Intelligence (UAI)*, 2006.

PATENTS

D. Krishnaswamy, R. Daley, N. Verma. Delivery of targeted content related to a learned and predicted future behavior based on spatial, temporal, and user attributed and behavioral constraints. *US Patent Application # 20110282964*, May 2010.

D. Krishnaswamy, N. Verma, V. Bychkovsky. Method and system using keyword vectors and associated metrics for learning and prediction of user correlation of targeted content messages in a mobile environment. *US Patent Application # 20090125462*, May 2009.

D. Krishnaswamy, N. Verma, V. Bychkovsky. Method and system for keyword correlation in a mobile environment. *US Patent Application # 20090125517*, May 2009.

INVITED TALKS

Distance preserving embeddings for Riemannian manifolds

- Carnegie Mellon University, Machine Learning Department (host Aarti Singh)
- IBM Research, Almaden (host Ken Clarkson)
- University of Washington, Math Department (host Marina Meila)
- Yahoo Labs, Bangalore (host Dhruv Mahajan)

Sample complexity of metric learning

- IBM Research, Almaden (host Ken Clarkson)
- Machine Learning seminar, Janelia Research Campus, HHMI (host Kristin Branson)

A tutorial on metric learning with some recent advances

- Bay Area Machine Learning Group (host Tony Tran)

An introduction to statistical theory of learning

- Neurotheory seminar, Janelia Research Campus, HHMI (host Shaul Druckmann)