Paramveer Dhillon

1717 South Street
Philadelphia PA 19146, U.S.A
+1-215-588-9636
dhillon@cis.upenn.edu
http://www.pdhillon.com

Education

2010–Present **Ph.D. in Computer and Information Science**,
University of Pennsylvania, Philadelphia PA, U.S.A.

+ Advisors: Prof. Lyle Ungar and Prof. James Gee

2007–2009 **M.S. in Computer and Information Science**,
University of Pennsylvania, Philadelphia PA, U.S.A.

- GPA: 3.95/4.00
- M.S. Thesis: Transfer Learning Using Feature Selection

+ M.S. Thesis Committee: Prof. Lyle Ungar (CIS), Prof. Ben Taskar (CIS) and Prof. Dean Foster (Wharton)

2003–2007 **B.E. in Electronics and Electrical Communication Engineering**,
Punjab Engineering College, Chandigarh, India.

- GPA: (According to Indian/British System): 76.98% (Honors with Distinction)

Research Interests

Machine Learning in general and its applications to NLP (Natural Language Processing), Web Data and Biomedical Imaging.

Publications

Journals (Heavily Refereed.)

1. "Minimum Description Length Penalization for Group and Multi-Task Sparse Learning",
Paramveer Dhillon, Dean Foster and Lyle Ungar.
Journal of Machine Learning Research (JMLR) 12, Feb. 2011.
(Highest Impact Factor Machine Learning Journal.)

Conferences (Heavily Refereed.)

13. "Two Step CCA: A new spectral method for estimating vector models of words",
Paramveer Dhillon, Jordan Rodu, Dean Foster and Lyle Ungar.
International Conference on Machine Learning (ICML 2012)
(Acceptance Rate: 27.3%)

12. "Spectral Dependency Parsing with Latent Variables",
Paramveer Dhillon, Jordan Rodu, Michael Collins, Dean Foster and Lyle Ungar.
Joint International Conference on Empirical Methods in Natural Language Processing and Conference on Natural Language Learning (EMNLP-CoNLL 2012)
(Acceptance Rate: 25.0%)

11. "Eigenanatomy improves detection power for longitudinal cortical change",
Brian Avants, **Paramveer Dhillon**, Benjamin Kandel, Corey McMillan, Murray Grossman, James Gee and Lyle Ungar.
International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2012)
(Acceptance Rate: 31.8%)

10. "Partial Sparse Canonical Correlation Analysis (PSCCA) for population studies in Medical Imaging",
Paramveer Dhillon, Brian Avants, Lyle Ungar and James Gee.
IEEE International Symposium on Biomedical Imaging (ISBI 2012)
(Acceptance Rate: 39.8%)

9. "Deterministic Annealing for Semi-Supervised Structured Output Learning",
Paramveer Dhillon, Sathiya Keerthi, Olivier Chapelle, Kedar Bellare and S. Sundararajan.
International Conference on AI and Statistics (AISTATS 2012)
(Acceptance Rate: 33.5%)

8. "Multi View Learning of Word Embeddings via Canonical Correlation Analysis",
Paramveer Dhillon, Dean Foster and Lyle Ungar.
Advances in Neural Information Processing Systems 24, (NIPS 2011)
(Acceptance Rate: 21.8%)

7. "Semi-supervised Multi-task Learning of Structured Prediction Models for Web Information Extraction",
Paramveer Dhillon, S. Sundararajan and Sathiya Keerthi.
International Conference on Information and Knowledge Management (CIKM 2011)
(Acceptance Rate: 15.0%)

6. "A New Approach to Lexical Disambiguation of Arabic Text",
Rushin Shah, **Paramveer Dhillon**, Mark Liberman, Dean Foster, Mohamed Maamouri and Lyle Ungar.
International Conference on Empirical Methods in Natural Language Processing (EMNLP 2010)
(Acceptance Rate: 25.0%)

5. "Feature Selection using Multiple Streams",
Paramveer Dhillon, Dean Foster and Lyle Ungar.
International Conference on AI and Statistics (AISTATS 2010)
(Acceptance Rate: 40.6%)

4. "Learning Better Data Representation using Inference-Driven Metric Learning (IDML)",
Paramveer Dhillon, Partha Pratim Talukdar and Koby Crammer.
Association of Computational Linguistics (ACL 2010)
(Acceptance Rate: 22.0%)

3. "Transfer Learning, Feature Selection and Word Sense Disambiguation",
Paramveer Dhillon and Lyle Ungar.
Association of Computational Linguistics (ACL-IJCNLP 2009)
(Acceptance Rate: 24.6%)

2. "Multi-Task Feature Selection Using the Multiple Inclusion Criterion (MIC)",
Paramveer Dhillon, Brian Tomasik, Dean Foster and Lyle Ungar.
European Conference on Machine Learning (ECML-PKDD 2009)
(Acceptance Rate: 24.9%)

1. "Efficient Feature Selection in the Presence of Multiple Feature Classes",
Paramveer Dhillon, Dean Foster and Lyle Ungar.
IEEE- International Conference on Data Mining (ICDM 2008)
(Acceptance Rate: 19.9%)

Workshops/Student Symposiums (Lightly Refereed.)

3. "Inference Driven Metric Learning for Graph Construction",
Paramveer Dhillon, Partha Pratim Talukdar and Koby Crammer.
North East Student Conference on AI (NESCAI 2010)

2. "Combining Appearance and Motion for Human Action Classification in Videos",
Paramveer Dhillon, Sebastian Nowozin and Christoph Lampert.
International Workshop on Visual Scene Understanding ViSU at CVPR 2009

1. "Robust Real-Time Face Tracking using an Active Camera",
Paramveer Dhillon.
CISIS Workshop, Lecture Notes in Computer Science (LNCS) 2009

Abstracts (Not formally published but have a oral/poster presentation.)

4. "Efficient Learning of Word Embeddings via Canonical Correlation Analysis" (Oral Presentation),
Paramveer Dhillon, Dean Foster and Lyle Ungar.
New York Academy of Sciences Machine Learning Symposium (NYAS-ML 2011)

3. "Improved Semi-Supervised Learning using Constraints and Relaxed Entropy Regularization via Deterministic Annealing" (Poster Presentation),
Paramveer Dhillon, Sathiya Keerthi, Olivier Chapelle, Kedar Bellare and S. Sundararajan.
New York Academy of Sciences Machine Learning Symposium (NYAS-ML 2011)

2. "Retinal nerve fiber thickness as a biomarker in Alzheimer's disease" (Oral Presentation),
Philip Cook, **Paramveer Dhillon**, Brian Avants, Laura Balcer, Murray Grossman, F. Swenson, James Gee.
Society of Neuroscience (SfN) Abstracts 2011

1. "White matter microstructure associated with resilience to social stress in rats" (Oral Presentation),
Philip Cook, **Paramveer Dhillon**, Brian Avants, Caroline Brun, S. Luz, Seema Bhatnagar, James Gee.
Society of Neuroscience (SfN) Abstracts 2011

Summer Research Internships

Summer 2011 **Yahoo! Research (Machine Learning Group)**, *Santa Clara, CA, U.S.A.*
Summer Intern (Worked with Sathiya Keerthi, Olivier Chapelle and Kedar Bellare)

- Worked on developing constraint driven sequence learning models for Information Extraction.
- The work resulted in publication at AISTATS 2012.

Summer 2010 **Yahoo! Research**, *Bangalore, India*.
Summer Intern (Worked with S. Sundararajan (Bangalore, India) and Sathiya Keerthi (Santa Clara, CA, U.S.A))

- Worked on Semi-Supervised and Multi-Task Learning of Structured Prediction Models.
- Developed novel Structured Prediction model for Information Extraction from totally unlabeled sites.
- The work resulted in publication at CIKM 2011.

Summer 2009 **Information Sciences Institute (ISI/USC)**, *Los Angeles CA, U.S.A.*
Summer Intern (Worked with Prof. David Chiang and Prof. Kevin Knight)

- Worked on Target Language Simplification for Machine Translation Systems.
- Worked on unsupervised log-linear models for Word Alignment.

Summer 2008 **Max Planck Institute for Biological Cybernetics**, *Tübingen, Germany*.
Summer Intern (Worked with Prof. Christoph Lampert and Prof. Bernhard Schölkopf)

- Worked on developing robust descriptors for Human Action Classification In Videos.
- The work resulted in publication at ViSU-CVPR 2009.

Summer 2006 **Computer Vision Center, Universitat Autónoma de Barcelona**, *Barcelona, Spain*.
Summer Intern (Under supervision of Dr. Jordi González and Dr. J. J. Villanueva)

- Developed algorithms for Detection and Tracking of Facial Features in image sequences and then implemented the algorithms in real- time on a distributed array of PTZ cameras and wrote code for Cooperation of PTZ (Pan Tilt Zoom) sensors in response to unambiguous segmentation.
- The work resulted in a CISIS Workshop paper.

Talks (Excluding Conferences.)

Nov. 4, 2011 **Temple University**, *NLP Group*, Philadelphia, PA, U.S.A.

July 12, 2011 **Yahoo! Research**, *Machine Learning Group*, Santa Clara, CA, U.S.A.

Nov. 21, 2010 **Machine Learning Lunch, CIS Department**, *Seminar Talk*, University of Pennsylvania, Philadelphia, PA, U.S.A.

Aug. 25, 2010 **Yahoo! Research**, *Search Sciences Group*, Bangalore, India.

Oct. 9, 2009 **SMiLe (Statistical Machine Learning and its Applications) Workshop**, *Invited Talk*, IBM T.J. Watson Research Center, Yorktown Heights, NY, U.S.A.

July 17, 2009 **Natural Language Seminar**, *Seminar Talk*, Information Sciences Institute (ISI)/USC, Los Angeles, CA, U.S.A.

Dec. 4, 2008 **Machine Learning Lunch, CIS Department**, *Seminar Talk*, University of Pennsylvania, Philadelphia, PA, U.S.A.

Aug. 7, 2008 **Department of Empirical Inference, (Tea Talk)**, *Seminar Talk*, Max Planck Institute for Biological Cybernetics, Tübingen, Germany.

Manuscript Reviewing Independent (Primary) Reviewer.

Conferences/Journals

2011 AISTATS, JMLR.

2010 IEEE TKDE, Radiology Journal.

Teaching Experience

Fall 2008 **Teaching Assistant**, *CIS 520 - Introduction to Machine Learning*.
Computer and Information Science, University of Pennsylvania
Instructor: Prof. Ben Taskar
- Duties: Holding office hours twice a week, holding biweekly recitation, grading homeworks and exams, formulating questions for midterm and final exam.
- CIS 520 constitutes a "core course" requirement for 1^{st} year CIS Ph.D students.

Spring 2011 **Teaching Assistant**, *CIS 320 - Introduction to Algorithms*.
Computer and Information Science, University of Pennsylvania
Instructor: Prof. Sanjeev Khanna
- Duties: Holding office hours twice a week, grading homeworks and exams, formulating questions for midterm and final exam.
- CIS 320 constitutes a "core course" requirement for CIS undergraduate students.

Spring 2009 **Teaching Assistant**, *CIT 595 - Computer Systems II*.
Computer and Information Science, University of Pennsylvania
Instructor: Diana Palsetia
- Duties: Holding office hours twice a week, grading homeworks and quizzes, making questions for quizzes.
- CIT 595 constitutes a "core course" requirement for 1^{st} year MCIT students.

Fall 2009 **Teaching Assistant**, *CIT 593 - Computer Systems I*.
Computer and Information Science, University of Pennsylvania
Instructor: Diana Palsetia
- Duties: Holding office hours twice a week, grading homeworks and quizzes, making questions for quizzes.
- CIT 593 constitutes a "core course" requirement for 1^{st} year MCIT students.

Spring 2010 **Teaching Assistant**, *CIT 595 - Computer Systems II*.
Computer and Information Science, University of Pennsylvania
Instructor: Diana Palsetia
- Duties: Holding office hours twice a week, grading homeworks and quizzes, making questions for quizzes.
- CIT 595 constitutes a "core course" requirement for 1^{st} year MCIT students.

Language Skills

Human

English, German, Indian (Punjabi and Hindi).

Computer

Programming	Java, Python, MATLAB.
Scripting/Markup	(D)HTML, LaTeX.

Awards

+ Received the highly prestigious and competitive **Provost's Fellowship** to pursue graduate studies (Ph.D) at University of Southern California (USC) *(Declined)*
+ Received Student Travel Award and volunteer scholarship for presenting the paper at ICML 2012.
+ Received Student Travel Award for presenting the paper at NIPS 2011.
+ Received NSF (National Science Foundation) Travel Award for presenting the paper at ICDM 2008.
+ Invited to present my research work at **SMiLe (Statistical Machine Learning and its Applications)** Workshop at IBM T.J Watson Research Center, Yorktown Heights, NY, U.S.A
+ Biography listed in 2010 Marquis Who's Who in Science and Engineering
+ Departmental Honors for outstanding performance in undergraduate studies.
+ College Color (a medal) for outstanding performance in extra-curricular activities in undergraduate studies.

References

Available on Request.