Sumit Chopra

Courant Institute of Mathematical Sciences
New York University

719 Broadway
Room No. 1204
New York, NY - 10003

917-945-2221
sumit@cs.nyu.edu
www.cs.nyu.edu/~sumit

RESEARCH INTERESTS

I am interested in Machine Learning and Pattern Recognition. I have focused on designing efficient inference and learning algorithms for models that can handle the uncertainties and interdependencies among samples in large scale data sets. To this end, I have worked on a diverse set of models, including Energy-Based Models, Graphical Models, Deep Learning Architectures, and Relational Graphical Models. These models have been applied in a variety of domains, such as, in economics for predicting the prices of real estate properties, in computer vision for object recognition, face verification, distance metric learning, manifold learning, and data visualization and in robotics for autonomous navigation.

EDUCATION

Ph.D in Computer Science Fall 2003 - Present*
Courant Institute of Mathematical Sciences, New York University, New York.
* Expected graduation date: Spring 2008.

Bachelor of Information Sciences Fall 1999 - Spring 2003
Computer Science Department, HansRaj College, University of Delhi, New Delhi.

PUBLICATIONS

Factor Graphs for Relational Regression. Sumit Chopra, Trivikraman Thampy, John Leahy, Andrew Caplin, and Yann LeCun. *Technical Report, TR2007-906, January 2007.*

Discovering the Hidden Structure of House Prices with a Non-Parametric Latent Manifold Model. Sumit Chopra, Trivikraman Thampy, John Leahy, Andrew Caplin, and Yann LeCun. *13th International Conference on Knowledge Discovery and Data Mining (KDD), San Jose CA, August 2007.*

Energy-Based Models in Document Recognition and Computer Vision. Yann LeCun, Sumit Chopra, Marc'Aurelio Ranzato, and Fu Jie Huang. *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR 2007).*

A Unified Energy Based Framework for Unsupervised Learning. Marc'Aurelio Ranzato, T-Lan Boureau, Sumit Chopra, and Yann LeCun. *Proceedings of the 2007 Conference on Artificial Intelligence and Statistics (AISTATS 2007).*

Output-Sensitive Algorithms for Optimally Constructing Upper Envelope of Straight Line Segments in Parallel. Neelima Gupta, and Sumit Chopra. *Journal of Parallel and Distributed Computing (2007).*

On-line Learning for Offroad Robots: Using Spatial Label Propagation to Learn Long-Range Traversability. R. Hadsell, P. Sermanet, J. Ben, A. Erkan, J. Han, M. Grimes, S. Chopra,

Y. Sulsky, B. Flepp, U. Muller, and Y. LeCun. *Technical Report, CBLL-TR-2007-01-01, January 2007.*

A Tutorial on Energy-Based Learning. Yann LeCun, Sumit Chopra, Raia Hadsell, Marc'Aurelio Ranzato, and Fu Jie Huang. *Predicting Structured Outputs, Bakir et al. (eds), MIT Press 2006.*

Efficient Learning of Sparse Overcomplete Representations with Energy-Based Model. Marc'Aurelio Ranzato, Christopher Poultney, Sumit Chopra, and Yann LeCun. *Advances in Neural Information Processing Systems 19, in Scholkopf et al. (eds), MIT Press, Cambridge, MA, 2006.*

Dimensionality Reduction by Learning an Invariant Mapping. Raia Hadsell, Sumit Chopra, and Yann LeCun. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), New York City, NY, July 2006.*

Learning a Similarity Metric Discriminatively, with Application to Face Verification. Sumit Chopra, Raia Hadsell, and Yann LeCun. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Diego, CA, June 2005.*

An Experimental Study of Different Strategies for DNS-Based Load Balancing. A. Agarwal, T. Agarwal, S. Chopra, A. Feldmann, N. Kammenhuber, P. Krysta, and B. Vocking. *In Proceedings of Euro-Par, Klagenfurt, Austria, August 2003.*

Optimal, Output-Sensitive Algorithms for Constructing Upper Envelope of Line Segments in Parallel. Neelima Gupta, Sumit Chopra, and Sandeep Sen. *In Proceedings of FSTTCS, Bangalore, India, December 2001.*

WORK/RESEARCH EXPERIENCE

Factor Graphs for Relational Regression Fall 2006 - Present
Research Assistant, NYU
Advisors: Prof. Yann LeCun, Prof. Andrew Caplin, and Prof. John Leahy
In many learning problems, the I.I.D assumption on input samples does not always hold; samples may be related to each other in ways so that the variables associated with one sample depend on the values of variables of related samples. Moreover, in many situations this dependence is hidden and cannot be measured quantitatively. We are working on designing a latent relational factor graph model to do prediction in such relational problems. We focus on problems involving continuous variables: *Relational Regression*. We have applied this framework for predicting the prices of real estate properties, which is inherently a relational problem.

Energy Based Models Fall 2005 - Present
Advisor: Prof. Yann LeCun
Absence of normalization in Energy based models (EBMs) requires them to be trained discriminatively. This is achieved by minimizing a contrastive loss function. However, not all the loss functions are capable of successfully training an EBM. We are working on designing sufficient conditions on loss functions whose minimization will result in a successful training in both supervised and unsupervised settings. In addition we are also working on designing an EM type training algorithm for EBMs involving latent variable – a deterministic EM. We are also working with advanced connectionist architectures, such as Autoencoders, Convolutional Networks, and Deep Belief Networks to design unsupervised learning algorithms for generating sparse representations of data. These algorithms have been applied to a variety of vision related tasks, such as handwritten digit recognition (for which we hold the current record) and image denoising.

Rigorous Early Termination Summer 2006
Summer Intern, NEC Research Laboratories, Princeton, USA
Mentor: Leon Bottou
A number of learning algorithms involve optimizing a penalized empirical risk. However since the objective function is already an approximation of the expected risk, there is no point in optimizing it all the way. We explored the theoretical and algorithmic consequences of early termination of such iterative optimization algorithms.

Learning Applied to Ground Robots (LAGR) Summer 2005
Research Assistant, NYU
Advisor: Prof. Yann LeCun
Worked on the design and implementation of an autonomous navigational system for a ground robot provided by DARPA. Here the challenge is to design the scheme based only on the standard hardware that is provided by them.

Learning an Invariant Mapping for Dimensionality Reduction Spring 2005 - Fall 2005
Advisor: Prof. Yann LeCun
Developed a technique for dimensionality reduction which is independent of any simple distance measure in the input space. This was done by discriminatively learning an invariant mapping from input to output space.

Learning Similarity Metric Discriminatively Summer 2004 - Fall 2004
Advisor: Prof. Yann LeCun
Developed a technique to learn a similarity metric - a function that maps input patterns into a target space such that the norm in the target space approximates the "semantic" distance in the input space. This similarity metric is learnt discriminatively using a Siamese network and applied to the problem of face verification.

Channel Assignment in Optical Networks Supporting Dense Wavelength Division Multiplexing (DWDM) Fall 2002 - Summer 2003
Advisors: Dr. Neelima Gupta and Prof. S. N. Maheshwari
The goal is to assign a physical path and a dedicated wavelength to a source-destination pair so as to maximize the number of requests that can be successfully processed. Developed algorithms with improved approximation factors for bidirected binary trees for specific instances (fully loaded leaf-to-leaf) and faster and more efficient algorithms for general instances.

k-Splittable Traffic Assignment on Identical Servers Summer 2002
Summer Intern, Algorithms and Complexity group (AG-1), Max Plank Institut für Informatik (MPI), Germany
Mentors: Prof. Berthold Vöcking, and Dr. Piotr Krysta
The problem is to assign a set of n traffic streams with different rates to a set of m identical servers with the objective of minimizing the maximum load over all servers. Each stream can be split at most k times. Designed and simulated a more efficient fixed parameter tractable algorithm over the recently presented one.

Computing the Upper Envelope of Line Segments in Parallel Fall 2001 - Spring 2002
Advisors: Dr. Neelima Gupta and Prof. Sandeep Sen
Proposed two randomized and one deterministic output size sensitive parallel algorithms for the problem of computing the upper envelope of n straight line segments in a plane. The algorithms are superior in the sense that they speed-up optimally (i.e. processor time product matches the lower bound) with the output size in the sub-logarithmic time domain.

TEACHING EXPERIENCE

Learning Robots : under-graduate course at NYU	Spring 2006
Foundations of Machine Learning : graduate course at NYU	Spring 2005
Machine Learning and Pattern Recognition : graduate course at NYU	Fall 2004
Distributed Programming : graduate course at NYU	Spring 2004 and Fall 2005
Fundamental Algorithms : graduate course at NYU	Fall 2003

OTHER SKILLS

Course Work
Foundations of Machine Learning, Machine Learning and Pattern Recognition, Advanced topics in Machine Learning, Non-Linear Optimization, Convex and Non-Smooth Optimization, Numerical Methods - 1, Computer Vision, Honors Programming Languages, Honors Algorithms.

Languages and Software
C, C++, Lisp, Lush (an object oriented lisp interpretor), MATLAB.

HONORS

Recipient of MacCraken fellowship award from New York University.
Ranked first in the undergraduate class, HansRaj College, Delhi University.
Member of the Indian Association of Researchers in Computer Science (IARCS).

REFERENCES

Prof. Yann LeCun, New York University, 212-998-3283 (O), yann@cs.nyu.edu
Prof. Foster Provost, New York University, 212-998-0806 (O), fprovost@stern.nyu.edu
Leon Bottou, NEC Research Laboratories, Princeton, 609-951-2732 (O), leon@bottou.org