

Navdeep Jaitly

511-35 Charles St. West,
Toronto, ON M4Y 1R6

ndjaitly@cs.toronto.edu

Ph: (647) 680-4744

Highlights of Qualifications:

- Played a pivotal role in the adoption of Deep Neural Networks for Speech Recognition at Google by implementing my pipeline into Google's systems during an internship. Trained the first few neural network models for the task at Google.
- Member of winning team in Kaggle Competition for Quantitative Structure Analysis Relationship (QSAR) sponsored by Merck.
- Exceptional problem-solving skills and success in national mathematical competitions (63rd in Putnam Mathematical Competition, 1996; top 31 in Indian National Mathematics Olympiad, 1993)
- Master's in Computer Science, University of Waterloo 2000;
- Finishing Ph.D. in Machine Learning under the supervision of Geoffrey Hinton at the University of Toronto.

PhD Intern, Google Inc., Mountain View (June '11 – Aug '11)

- Used Deep Neural Networks for acoustic modeling of Google's Voice Search and YouTube data. This project reduced the word error rate of Google's best Voice Search system from 16.0% to 11.8%. The YouTube system error also achieved new state of the art, improving the word error rate of the best system from 52.3% to 46.6% .
- Implemented the Deep Neural Networks / Hidden Markov Model system into the speech recognizer at Google.

PhD Student, Machine Learning, University of Toronto (Sept '08 – current)

- Finishing thesis on "Discovering features in speech signals using Deep Learning".
- Developed algorithms and approaches for modeling distributions of images and speech with Capsules and Templated Autoencoders.
- Developed a Bayesian method to infer 3D structure of proteins from electron cryomicroscopy tomograms.

Senior Research Scientist IV (Jan '08 – July'08) & Senior Research Scientist III (Sept '04 – Dec '07), Pacific Northwest National Laboratory

- Developed Models to statistically characterize noise and signal in high throughput scientific data and used the model to develop a search engine which allowed optimal selection of results at different levels of confidence.
- Developed algorithms for discovering common features in hundreds of mass spectrometry datasets using dynamic programming algorithms to align time series data and hierarchical clustering algorithms to group features.

- Developed improved signal processing and peak detection algorithms for mining of mass spectra where each dataset is of an average size of 8 Gb and sped up previous algorithms by 80%.
- Developed two analytical software tools: Decon2LS (open source) and MultiAlign in C++ and C# for manual and automated analysis of spectral data. Both tools were well used and continue to get a strong rate of citations.

Bioinformatics Lead, Collaborative Project with Université De Montréal, Apr '04 – Sept '04

- Developed Peptide Miner software for feature detection in mass spectrometry data, in collaboration with Dr. Pierre Thibault, University of Montreal.
- Software developed is now used to perform LC-MS data analysis at the Institute of Cancerology and Immunovirology, University of Montreal.

Group Leader, Senior Bioinformatics Scientist (Mar'01-'02), Caprion Pharmaceuticals Inc., Mar '01 – Apr '04

- Investigated and implemented several Data Mining, and AI approaches to discover, classify and predict trends in Biochemical data sets from Mass Spectrometry.
- Developed *Deep Search*, an algorithm to identify genetic differences, mutations and Biochemical Modifications using peptide mass spectra.
- Developed clustering algorithm to organize and assess the certainty of Protein Identification. Used it for comparing results of experiments from different conditions as a first pass at detecting biomarkers.

Software Developer, SMART Research Group, IBM Toronto Labs, Aug '00 – Mar '01

- Modernized existing tool (Smart Guide) for automatic database configuration of DB2, IBM's Universal Database.
- Laid a framework for learning of optimal database configurations in DB2.
- Project chosen as Showcase of DB2's Technology in Annual DB2 conference in Germany.

Education:

- PhD in Computer Science with focus on Machine Learning. University of Toronto. (5th year)
- Master's in Computer Science (MMath). University of Waterloo, Waterloo, ON. August 2000
- B.A., Computer Science & Mathematics. Hanover College, Hanover, IN. May 1998.
- B.Tech (one year) Indian Institute of Technology, I.I.T. Delhi.

Publications & Patents:

- Graves, A. Jaitly, N. Mohamed, A.R. *Automatic Speech Recognition and Understanding Workshop*. 2013. Hybrid Speech Recognition with Deep Bidirectional LSTM.
- Jaitly, N. and Hinton, G.E. *International Conference on Machine Learning (ICML)*. 2013. Deep Learning for Audio, Speech and Language Processing Workshop. Vocal Tract Length Perturbation (VTLP) improves speech recognition.

- LaMarche, B.L, Crowell, K.L., Jaitly, N. et. al. *BMC Bioinformatics*. 14(1):49. MultiAlign: a multiple LC-MS analysis tool for targeted omics analysis.
- Jaitly, N. and Hinton, G.E. *Interspeech*. 2013. Using an autoencoder with deformable templates to discover features for automated speech recognition.
- Hinton, G.E., et. al. *Signal Processing Magazine, I.E.E.E.* 29(6):82-97. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups.
- Jaitly, N., Nguyen, P., Senior, A.W. and Vanhoucke, V. *Interspeech*. 2012. Application of Pretrained Deep Neural Networks to Large Vocabulary Speech Recognition.
- Jaitly, N. and Hinton, G. E. *Advances in Neural Information Processing Systems 24*, Deep Learning workshop, Grenada, Spain. A new way to learn acoustic events.
- Wallach I., Jaitly N., Nguyen K., Schapira M., Lilien R. *Journal of Chemical Information Modelling*. 2011 Aug 22;51(8):1817-30. Normalizing molecular docking rankings using virtually generated decoys.
- Jaitly, N., Hinton, G.E. *ICASSP 2011. Special session: Innovative Representations of Audio*. Learning a better representation of speech sound waves using Restricted Boltzmann Machines.
- Jaitly, N., Brubaker, M.A., Rubinstein, J.L., Lilien, R.H. *Bioinformatics*. 2010. Oct; 26(19):2406-15. A Bayesian method for 3D macromolecular structure inference using class averages from single particle electron microscopy.
- Jaitly, N., Brubaker, M.A., Rubinstein, J.L., Lilien, R.H. *Machine Learning in Computational Biology (MLCB), NIPS 2009 workshop*. A Bayesian method for 3D reconstruction of macromolecular structure using class averages from single particle electron microscopy.
- Wallach, I., Jaitly, N., Lilien, R.H. *PLoS One*. 2010. Aug; 5(8):e12063. A Structure-Based Approach for Mapping Adverse Drug Reactions to the Perturbation of Underlying Biological Pathways.
- Umar A, Kang H, Timmermans AM, Look MP, Meijer-van Gelder ME, den Bakker MA, Jaitly N, Martens JW, Luidert TM, Foekens JA, Pasa-Tolić L. *Molecular and Cellular Proteomics*. 2009. Jun;8(6):1278-94. Identification of a putative protein profile associated with tamoxifen therapy resistance in breast cancer.
- Jaitly N, Mayampurath A, Littlefield K, Adkins JN, Anderson GA, Smith RD. *BMC Bioinformatics*. 2009. Mar 17;10:87. Decon2LS: An open-source software package for automated processing and visualization of high resolution mass spectrometry data.
- Ansong C, Yoon H, Porwollik S, Mottaz-Brewer H, Petritis BO, Jaitly N, Adkins JN, McClelland M, Heffron F, Smith RD. *PLoS One*. 2009;4(3):e4809. Global systems-level analysis of Hfq and SmpB deletion mutants in Salmonella: implications for virulence and global protein translation.
- Qian WJ, et. al. *Journal of Proteome Research* 2009 Jan;8(1):290-0. Large-scale multiplexed quantitative discovery proteomics enabled by the use of an (18)O-labeled "universal" reference sample.
- Tolmachev AV, Monroe ME, Purvine SO, Moore RJ, Jaitly N, Adkins JN, Anderson GA, Smith RD. *Analytical Chemistry*. 2008 Nov 15;80(22):8514-25. Characterization of strategies for obtaining confident identification in bottom-up proteomics measurements using hybrid FTMS instruments.

- Sorensen, C.M., Ding, J., Jaitly, N., et. al. *Journal of Chromatography B Analytical Technologies in the Biomedical and Life Sciences*. 2008 May 7. Application of the accurate mass and time tag approach in studies of the human blood lipidome.
- Polpitiya A.D., Qian, W.J., Jaitly N., et. al. *Bioinformatics*. 2008 Jul 1;24(13):1556-8. DANTE: A statistical tool for quantitative analysis of -omics data.
- Ham B.M., et. al. *Journal of Proteome Research*. 2008 Jun; 7(6):2215-21. The influence of sample preparation and replicate analyses on HeLa Cell phosphoproteome coverage.
- Mayampurath A.M., Jaitly N., Purvine S.O., Monroe M.M., Auberry K.A., Adkins J.N., Smith R.D. *Bioinformatics*. 2008 Apr 1;24(7):1021-3. DeconMSn: A software tool for accurate parent ion monoisotopic mass determination for tandem mass spectra.
- Petyuk, V.A., Jaitly, et. Al. *Analytical Chemistry* 2008 Feb 1;80(3):693-706. Elimination of systematic mass measurement errors in liquid chromatography-mass spectrometry based proteomics using regression models and a priori partial knowledge of the sample content.
- Yang, F., Jaitly, N., Jayachandran, H., et. Al. *Journal of Proteome Reseach* 2007 Nov;6(11):4489-97. Epub 2007 Oct 12. Applying a Targeted Label-Free Approach Using LC-MS AMT Tags to Evaluate Changes in Protein Phosphorylation Following Phosphatase Inhibition.
- Gupta, N., Tanner, S., Jaitly, N., Adkins, J.N., Liptono, M., Edwards, R., Romine, M., Osterman, A., Bafna, V., Smith, R.D., Pevzner, P.A. *Genome Research*. 2007 Sep; 17(9): 1362-77. Whole proteome analysis of post-translational modifications: Applications of mass-spectrometry for proteogenomic annotation.
- Liu, T., Belov, M.E., Jaitly, N., Qian, W.J., Smith, R.D. *Chemical Reviews* 2007 Aug; 107(8): 3621-53. Accurate Mass Measurements in Proteomics.
- Ding, J., Sorensen, C.M., Zhang, Q., Jiang, H., Jaitly, N., Livesay, E.A., Shen, Y., Smith, R.D., Metz, T.O. *Analytical Chemistry* 2007 Aug 15; 79(16): 6081-93. Capillary LC coupled with high-mass measurement accuracy mass spectrometry for metabolic profiling.
- Monroe, M.E., Tolic, N., Jaitly, N., Shaw, J.L., Adkins, J.N., Smith, R.D. *Bioinformatics* Aug 1;23(15):2021-3. VIPER: an advanced software package to support high-throughput LC-MS peptide identification.
- Jaitly, N., Monroe, M.E., Petyuk, V.A., Clauss, T.R.W, Adkins, J.N., Smith, R.D. *Analytical Chemistry* 2006 Nov 1;78(21):7397-409. Robust Algorithm for Alignment of Liquid Chromatography-Mass Spectrometry Analyses in an Accurate Mass and Time Tag Data Analysis Pipeline.
- Kearney, P. et. al. (*US Patent Application being filed 2006*). TAT-001 and Methods of Assessing and Treating Cancer.
- Sharma, S., Simpson, D.C., Jaitly, N., Mayampurath, A.M., Smith, R.D., Pasa-Tolic, L. *Journal of Proteome Research*. 2007 Feb; 6(2):602-10. Proteomic Profiling of Intact Proteins Using WAX-RPLC 2-D Separations and FTICR Mass Spectrometry
- Petyuk, V.A. et.al. *Genome Research*. 2007 Mar;17(3):328-36. Spatial mapping of protein abundances in the mouse brain by voxelation integrated with high-throughput liquid chromatography-mass spectrometry.
- Tolmachev A.V., Monroe M.E., Jaitly N., Petyuk V.A., Adkins J.N., Smith R.D. *Analytical Chemistry*. 2006 Dec 15;78(24):8374-85. Mass measurement accuracy in analyses of highly complex mixtures based upon multidimensional recalibration

- Norbeck, A.D., Callister, S.J., Monroe, M.E., Jaitly, N., Elias, D.A., Lipton, M.S., Smith, R.D. *Journal of Microbiological Methods*. 2006 Dec;67(3):473-86. Proteomic approaches to bacterial differentiation.
- Kiebel GR, Auberry KJ, Jaitly N, Clark DA, Monroe ME, Peterson ES, Tolic N, Anderson GA, Smith RD. *Proteomics*. 2006 Mar;6(6):1783-90. PRISM: A Data Management System for High-Throughput Proteomics.
- McIntire, T.M., Lea, A.S., Gaspar, D.J., Jaitly, N., Dubowski, Y., Li, Q., Finlayson-Pitts, B.J. *Phys Chem Chem Phys*. 2005 Oct 21;7(20):3605-9. Unusual Aggregates from the Oxidation of Alkene Self-Assembled Monolayers: A Previously Unrecognized Mechanism for SAM Ozonolysis?
- Kearney, P, Jaitly, D., Eng, K. (*US Patent Application# 20060122785*). Constellation Mapping: Software and Methods to match complex peptide maps using nonlinear alignment methods.
- Jaitly, D, Page-Belanger, R., Faubert, D, Thibault, P, Kearney, P. MSMS Peak Identification through Maximum Likelihood Methods And its Application to Analysis of Tandem Mass Spectrometry data. *ISMB 2004*.
- Swamy, S, Furtos-Matei, A, Jaitly, N, Kearney, P, Thibault, P. GlycoTrack: Tool for Automated N-Linked Glycopeptide Identification and Analysis. Poster, *PSB 2003 (Also, US Patent Application# 20040248317)*.
- Jaitly, D, Kearney, P, Lin, G and Ma, B. Methods for Reconstructing the History of Tandem Repeats and Their Application to the Human Genome. *Journal of Computer and System Sciences* 65(3): 494-507. 2002.

Awards and honors:

- Member of the winning team in the Kaggle Competition for Quantitative Structure Analysis Relationships (QSAR) sponsored by Merck, 2012.
- Helen Sawyer Hogg Graduate Admission Award, University of Toronto, 2008.
- Nominated for Environmental and Molecular Sciences Laboratory (EMSL) **Director's Award** for “.contributions in support of Environmental and Molecular Sciences Laboratory's strategy for achieving high-impact science and marquee capabilities..”, Pacific Northwest National Laboratory, November 2006.
- **Outstanding Performance Award**, Pacific Northwest National Laboratory, 2006 for Development and Implementation of Advanced Algorithms for Mass and Time Alignment of Proteomics Data
- Outstanding Performance Award, Pacific Northwest National Laboratory, 2005.
- Ranked 63rd in **William Lowell Putnam Mathematics Competition**, Dec 1996.
- Ranked in the top 31 in the **Indian National Mathematics Olympiads**, 1993.
- Dean's List