Avneesh Singh Saluja

CONTACT INFORMATION

45A Pearl Street
San Francisco, CA
94103 USA

☎: +1-650-526-8792
✉: avneesh@cs.cmu.edu
: http://www.cs.cmu.edu/~avneesh

RESEARCH INTERESTS

Probabilistic models for natural language processing; statistical machine translation; distributed data processing; machine learning.

EDUCATION

Carnegie Mellon University, Pittsburgh, PA

Ph.D., Electrical & Computer Engineering, September 2015

- Dissertation: "Low-dimensional Context-dependent Translation Models"
- Advisors: Chris Dyer & Ian Lane
- Affiliated with Language Technologies Institute, School of Computer Science
- Recipient, Bertucci Fellowship (College of Engineering), 2015
- Recipient, eBay Graduate Fellowship, 2014

Carnegie Mellon University, Pittsburgh, PA

M.S. in Electrical & Computer Engineering, May 2013

- Graduate coursework in statistics, machine learning, statistical learning theory, speech recognition, probabilistic graphical models, structured prediction, and optimization.

Stanford University, Stanford, CA

B.S. *with Distinction* in Electrical Engineering, April 2007

- Signal processing concentration
- President's Award for Academic Excellence, Tau Beta Pi

PROFESSIONAL EXPERIENCE

10/2015–present **Airbnb**
San Francisco, CA
Machine Learning Scientist
Driving multiple machine learning initiatives within product groups across the company, focusing on user-generated text and images

6/2014–8/2014 **eBay Inc.**
San Jose, CA
Research Intern, MT Group
Project: modeling compositional and non-compositional semantics, with applications in machine translation

5/2013–8/2013 **Microsoft Research**
Redmond, WA
Research Intern, Statistical MT Group
Project: graph-based semi-supervised learning for phrase table expansion

5/2012–8/2012 **IBM Research T.J. Watson Center**
Yorktown Heights, NY
Research Intern, Statistical NLP Group
Project: graph-based unsupervised word similarities from multiple feature types

7/2009–5/2010 **July Systems**
Bangalore, India
Lead Manager, Product & Product Marketing
Product managed won *Best Mobile International* category at 2010 Mobile Excellence Awards

7/2007–3/2009 **Goldman Sachs**
New York, NY & Hong Kong
Financial Analyst, Structured Equity Solutions Group
Series 7, 55, and 63 certified (US NASD), Paper 1 certified (HK SFC)

Publications

Journals

A. Saluja and Y. Zhang; Online Discriminative Learning for Machine Translation with Binary-valued Feedback. *Machine Translation*, Vol. 28 (2) pp. 69-90, 2014.

Conferences

A. Saluja, C. Dyer, and S. Cohen; Latent Variable Synchronous CFGs for Hierarchical Translation. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar. October 2014.

A.P. Parikh, **A. Saluja**, C. Dyer, and E.P. Xing; Language Modeling with Power Low Rank Ensembles. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar. October 2014.**(Best Paper Nominee)**

A. Saluja, H. Hassan, K. Toutanova, and C. Quirk; Graph-based Semi-Supervised Learning of Translation Models from Monolingual Data. In *Proceedings of the Association for Computational Linguistics (ACL)*, Baltimore, MD. June 2014.

A. Saluja, I. Lane, and Y. Zhang; Machine Translation with Binary Feedback: a Large-Margin Approach. In *Conference of the Association for Machine Translation in the Americas (AMTA)*, San Diego, CA. October 2012.

A. Saluja, P. Sundararajan, and O.J. Mengshoel; Age-Layered Expectation Maximization for Parameter Learning in Bayesian Networks. In *Artificial Intelligence & Statistics (AIStats)*, La Palma, Spain. April 2012.

A. Saluja, I. Lane, and Y. Zhang; Context-aware Language Modeling for Conversational Speech Translation. In *Machine Translation Summit XIII*, Xiamen, China. September 2011.

Workshops

A. Saluja and J. Navrátil; Graph-based Unsupervised Learning of Word Similarities Using Heterogeneous Feature Types. In *TextGraphs 8: Graph-based Methods for Natural Language Processing*, Seattle, WA. October 2013.

A. Saluja and B. Kveton; Semi-Supervised Learning with Cover Trees. In *Big Learning: 2011 NIPS Workshop on Parallel and Large-Scale Machine Learning*, Granada, Spain. December 2011.

A. Saluja, F. Mokaya, M. Phielipp, and B. Kveton; Automatic Identity Inference for Smart Televisions. In *AAAI 2011 Workshop on Lifelong Learning*, San Francisco, CA. August 2011.

Preprints & Technical Reports

A. Saluja, M. Pakdaman, D. Piao, and A.P. Parikh; Infinite Mixed Membership Matrix Factorization. *arXiv:1401:3413*. January 2014.

Invited Talks

June 2016	OpenAir Conference, San Francisco: *Machine Learning in a Community-driven Marketplace*
October 2016	Facebook, Menlo Park: *Neural Networks and NLP at Airbnb*
November 2016	ReWork Machine Intelligence Summit, New York: *Extracting Customer Insights at Airbnb*

Professional Service

Co-organizer	Modern Machine Learning and Natural Language Processing Workshop held at NIPS 2014
Program Committee	Multilingual and Crosslingual Methods in Natural Language Processing Workshop held at NAACL 2016
Reviewer	IWSLT 2013, NAACL 2016, ACL 2016, EACL 2017, JMLR

Teaching Experience

8/2014–12/2014	**18-751** Applied Stochastic Processes Teaching Assistant and Guest Lecturer
1/2013–5/2013	**18-799M** Advanced Machine Learning Teaching Assistant and Guest Lecturer

Languages

Programming:	Python, C++, C#, MATLAB/Octave, bash
Human:	English, Hindi, Urdu, Nepalese, Spanish (fluent); Punjabi (conversational); Arabic (basic)

Personal

Lived in 9 countries on 6 continents, with K-12 schooling in international schools
MENSA International member
Interests: violin, guitar, soccer, cricket, and squash

References

Chris Dyer, Assistant Professor, LTI (SCS), MLD (SCS), Carnegie Mellon University, `cdyer@cs.cmu.edu`

Ian Lane, Assistant Research Professor, ECE (CIT), LTI (SCS), CMU-SV, Carnegie Mellon University, `ianlane@cs.cmu.edu`

Joy (Ying) Zhang, Research Scientist, Facebook Inc., `joyzhang@fb.com`