Ashudeep Singh

Computer Science PhD Student, Cornell University

349 Gates Hall
Cornell University
Ithaca, NY, USA-14850
(607)-379-7806
ashudeep@cs.cornell.edu
www.ashudeepsingh.com

Education

2015-present **Ph.D. Computer Science**, *Cornell University*, Ithaca, NY.
Advisor: Thorsten Joachims
Relevant Courses: Machine Learning Theory, Advanced topics in Machine Learning, Design and Analysis of Algorithms
Grade Point Average: 4.00

2010–2015 **B.Tech.-M.Tech. Dual Degree**, *Indian Institute of Technology Kanpur*, India.
M.Tech. Cumulative Performance Index (CPI)– 10/10
B.Tech. Cumulative Performance Index (CPI)– 9.6/10

Awards and Achievements

2015 **Ranked first** in the M.Tech. batch of 108 students graduating in 2015.

2011,2012, 2013 Awarded **Academic Excellence Award** for outstanding academic achievements at IIT Kanpur.

2010–2014 Awarded **CBSE Merit Scholarship** for Professional Studies by Central Board of Secondary Education, India.

2012 Recipient of **Summer Undergraduate Research Grant for Excellence (SURGE)**, granted by Dean Resource Planning and Generation, IIT Kanpur.

2010 Awarded the **Certificate of Merit** in English for being in the top 0.1% students in the country for All India Senior School Certificate Examination.

2010 Represented Chandigarh region in the **INChO** (Indian National Chemistry Olympiad).

2009 Placed in **State-wide Top 1%** in National Standard Examination in Physics (NSEP 2009) conducted by IAPT (Indian Association of Physics Teachers, Pune, India).

Publications

Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, Thorsten Joachims. **"Recommendations as Treatments: Debiasing Learning and Evaluation"** In Proceedings of The International Conference on Machine Learning (ICML), 2016. [pdf]

David Adamson, Akash Bharadwaj, Ashudeep Singh, Colin Ashe, David Yaron, Carolyn P. Rosé. **"Predicting Student Learning from Conversational Cues"**. In Proceedings of 12th International Conference of Intelligent Tutoring Systems (ITS), Honolulu, HI, USA, June, 2014. [pdf]

David Adamson, Divyanshu Bhartiya, Biman Gujral, Radhika Kedia, Ashudeep Singh, Carolyn P. Rosé. **"Automatically Generating Discussion Questions"**. In Proceedings of 16th International Conference of Artificial Intelligence in Education (AIED), Memphis, TN, USA, July, 2013. [pdf]

Research Internships

May–July 2016 **Microsoft Research Lab**, *New York City, NY*.

"Contextual Bandits for Personalization in fitness tracking applications"

*Research internship project by **John Langford**(MSR NYC) and **Ryen White**(Microsoft Health Intelligence and MSR Redmond)*

The work was aimed to frame the task of providing reminders and notifications to users in Microsoft Health app service as a Contextual Bandits problem, so as to personalize these messages to optimize user's health and fitness.

- We used user's demographics and current health and fitness statistics as the context and calculate rewards based on the change in health and fitness measurements of the user.
- Worked with a state of the art decision service named Multi-world testing (MWT) to serve as the interface between the Microsoft Health and the app. The service is based on Vowpal Wabbit and returns the ranking of a set of messages to be shown to the user that maximizes the rewards.

May–July 2014 **Cornell University**, *Ithaca, NY*.

"Using Preference Data to Embed documents in Metric spaces"

*Research Project Mentored by **Prof. Thorsten Joachims**(Cornell University)*

The work is aimed at using human interaction signals to embed documents onto a low dimensional space. The data used is click-logs for user sessions on `arxiv.org`. We also try to embed the user sessions into the same space to facilitate its application in document recommendations and personalized search.

- Used **Logistic Markov Embedding** (LME) approach after decomposing user-sessions into first order markov chains. Utilised the feasible set information due to presentation and introduced a de-biasing feature vector to the model to **overcome the *bias*** because of presentation and browsing behavior.
- The embeddings perform **much better than traditional n-gram models** on predicting user clicks and also learn an intuitive low-dimensional space representation without using any text features.
- Currently working on **embedding user-sessions** onto the same metric space to represent user's intent and hence provide better recommendations and search results.

May–July 2013 **Carnegie Mellon University**, *Pittsburgh, PA*.

"A Computational Model for Quantitative Discourse Analysis in a Collaborative Learning Setting" [report]

*Research Project Mentored by **Prof. Carolyn P. Rosé** (Language Technologies Institute, CMU)*

We aimed to identify quantitative metrics that can be used to discriminate between successful and unsuccessful groups involved in a collaborative learning task, using text based chat transcripts.

- Used text features along with integer linear programming constraints to **disentangle the discourse** into partitioned sequences with annotations consisting of Knowledge-seeking and providing statements.
- Used standard **sequence modelling techniques over these annotated sequences** of dialogue at the discourse level along with text features to anticipate performance.
- Obtained **f-score of 0.728 for segmentation** of discourse. Also, obtained significant improvement in prediction accuracy using the structure of the dialogue at discourse level over simple text features.

Dec. 2012 **Internship Programme in Technology Supported Education**, *Winter School, Bangalore.*
"Question Generation for Discussion Facilitation" [report]
*Research Project Mentored by **Prof. Carolyn P. Rosé** (Language Technologies Institute, CMU)*
The motivation of the work was to encourage discussion and reasoning amongst students in a class through an intelligent tutoring system which generates questions that initiate discussion over a certain text.

- The first part involved **extraction of sentences** from a summary that are the best abstraction of the whole text using techniques like- LSA, Tf-Idf, Cosine Similarity and Jaccard Coefficient.
- For Question Generation, we modified a pre-existing implementation of question generation that replaces the **semantically labelled entities with WH-words**. For scoring and ranking questions, we use objectivity scores from SentiWordNet Corpus.

May–July 2012 **Summer Undergraduate Research Grant for Excellence (SURGE)**, *IIT Kanpur.*
"Logic Studio: Automatic Problem Generation in Propositional Logic" [report]
*Research Project Mentored by **Dr. Sumit Gulwani** (Microsoft Research, Redmond).*
The project comprised of generating deduction problems, hints and solutions to problems in Propositional Logic, which was part of a larger project which aims at building an Intelligent Tutor for Logic Course.

- Each proposition was represented as a **bit-vector of its truth values** to efficiently represent and search through the exponential sized problem space.
- The deduction process was **represented as a graph** with edges representing deduction using standard deduction formulas and equivalencies. Solutions are the paths through these graphs and new problems are generated from exploring nearby nodes.

Teaching

- **Teaching Assistant** for **CS4786: Machine Learning for Data Science** **(Spring 2016)**
  Faculty Instructor: Prof. Karthik Sridharan
  - Organized weekly office hours and problem-solving sessions.
  - Assisted the Faculty Instructor in designing the course content as well as problems for labs and exams.
- **Teaching Assistant** for **CS4780/5780 : Machine Learning** **(Fall 2015)**
  Faculty Instructor: Prof. Kilian Weinberger
  - Organized weekly office hours and regular problem-solving sessions.
  - Assisted the Faculty Instructor in designing the course content as well as problems for labs and exams.
- **Teaching Assistant** for **CS679–Machine Learning for Vision** **(Spring 2015)**
  Faculty Instructor: Prof. Vinay P. Namboodiri
- **Tutor (Graduate Student Instructor)** for **ESC101–Fundamentals of Computing** **(Fall 2014)**
  Faculty Instructor: Prof. Amey Karkare
  - Organized weekly tutorials and problem-solving sessions.
  - Assisted the Faculty Instructor in designing the course content as well as problems for labs and exams.

Skill Set

- **Programming Languages** – C, C++, Python, Java, C#, R, Lua
- **Web Development** – HTML, CSS, PHP, JavaScript
- **Other Tools** – Tensorflow, Torch, Shell Scripting, Matlab, SQL, Octave, LaTeX

Positions of Responsibility and Extra Curricular Activities

- Coordinated the PhD Visit Day 2016 for Cornell Computer Science Department as the *Visit Day Czar* with other czars, department staff and volunteers.
- **Student Guide, Counselling Service, IIT Kanpur** **(2011–12)**
  Guided freshmen for a year to adjust to the campus environment.
- **Academic Mentor, Counselling Service, IIT Kanpur** **(2011–12)**
  Conducted remedial classes for subjects like Fundamentals of Computing & Mathematics-I,II
- **Link Student, Counselling Service, IIT Kanpur** **(2012–13)**
  Responsible for helping two academically deficient students.

- **Member of Hall Executive Committee**, Hall 9, IIT Kanpur **(2011–12)**

- **Secretary, Hospitality Cell**, Techkriti 2011 **(2011)**.
- Co-Developed an Android app "MAP A FEST" that could display events currently going on during a campus festival or otherwise, allow users to update location and view friend's location on a Google Map, using Google Map API, mySQL databases and *facebook* APIs, during the Hack day by *Yahoo!* HACKU 2012.
- Placed 2^{nd} in **Madmen (Video Ad Making Competition)** in *Spectrum, IIT Kanpur Films and Media Festival, 2012*.