Ashudeep Singh

CONTACT INFORMATION

Final year Dual Degree Student
Department of Computer Science and Engineering
Indian Institute of Technology Kanpur

www.ashudeepsingh.com
e-mail: ashudeep@iitk.ac.in
Phone: +91 97941 40033

RESEARCH INTERESTS

Machine Learning and Statistical Natural Language Processing

EDUCATION

Indian Institute of Technology Kanpur, India

B.Tech.& M.Tech. Dual Degree in Computer Science and Engineering **(2010–present)**

- **Master of Technology** Cumulative Performance Index (CPI) of 10 on scale of 10
- **Bachelor of Technology** Cumulative Perfromance Index (CPI) of 9.6 on scale of 10

AWARDS AND ACHIEVEMENTS

- **Ranked first** in the M.Tech. batch of 108 students graduating in 2015.
- Awarded **Academic Excellence Award** for years 2010-11, 2011-12 and 2012-13 for outstanding academic achievements at IIT Kanpur.
- Awarded **CBSE Merit Scholarship** for Professional Studies for years 2010-14 by Central Board of Secondary Education, India.
- Recipient of **Summer Undergraduate Research Grant for Excellence (SURGE)** 2012, granted by Dean Resource Planning and Generation, IIT Kanpur.
- 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) 2010.
- 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

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

Cornell University **(May–July 2014)**

"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 performs much better than traditional n-gram models on predicting user clicks and also learns 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.

Carnegie Mellon University **(May–July 2013)**

"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 Mayfield et al.'s approach to disentangle the discourse into partitioned sequences with annotations consisting of Knowledge-seeking and Knowledge-providing statements. Used text features in supplement to standard sequence modelling techniques over these annotated sequences of dialogue at the discourse level.
- Showed that using the structure of the dialogue at discourse level significantly increases the prediction accuracy over simple text features.

Internship Programme in Technology Supported Education (Winter School, Bangalore) **(Dec 2012)**
"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.

Summer Undergraduate Research Grant for Excellence (SURGE) at IIT Kanpur **(May–July 2012)**
"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 formulas and equivalencies. Solutions were the paths through these graphs and new problems were generated from searching nearby nodes.

Selected Projects

Scene Recognition using mid-level CNN features **(Jan–Apr 2014)**
Computer Vision Course Project under Prof. Vinay P. Namboodiri [report]

- Computed 5^{th} and 6^{th} layer features from a pre-trained Conv-Net on ImageNet dataset using *Decaf* for the cells of the spatial pyramid representation for images from MIT-67 and 15-scene datasets.
- Compared classification results obtained for SVMs trained on 5^{th} and 6^{th} layer features.
- Experimented with a reconfigurable parts model representation of the images with relative configuration as the latent variable.

Recommendation System for movielens dataset **(Sept–Nov 2013)**
CS771 Course Project under Prof. Harish Karnick [report]

- Used SVD decomposition of preprocessed User-Movie Matrix to discover the latent features for users as well as movies and hence predict ratings of unseen movies for each user given their previous ratings, user demographic features and movie genre information.
- Pre-processing of User-Movie matrix was done by seeding the unknown ratings using k-NN and linear regression based prediction.
- Obtained RMSE comparable to the state-of-the-art for the $100k$ ratings *movielens* dataset.

Semantic approach to Summarization **(July–Nov 2013)**
CS498 Project under Prof. Harish Karnick [paper][code]

- Implemented a technique to identify clusters from text labelled according to FrameNet annotation, based on similarities between verb frames(verbs and arguments). Sentences were generated from the frames in a single cluster using a context free grammar based language generation technique.
- The summaries hence generated yielded good scores when evaluated against human summaries by human evaluators on 4 different aspects including Information content and Abstractness.

Student Response Analysis using Textual Entailment **(Sept–Nov 2013)**
Natural Language Processing Course Project under Prof. Amitabha Mukerjee [report][code]
The work involves a 5-way classification task, recognizing the extent of correctness of a student answer given a question and a few reference answers.

- We explored different text similarity and overlap features to train models on and also used various techniques used in Recognizing Textual Entailment Challenges since 2006.

Motion Tracking using Occlusion States **(Feb–Apr 2013)**
Artificial Intelligence Course Project under Prof. Amitabha Mukerjee [report]

- Formalized a transition graph for a set of 14 states that define occlusions in multi-object systems.
- Theoritically proved the validity of transitions using the logical formulations as given in the definitions and verified the transitions on real-world visual scenes and also tried mining events from a real world scene using signature transition sequences.

Visit ashudeepsingh.com/projects.html for a full list of projects and term-papers.

Relevant Courses Taken

- **ML and AI**: Machine Learning, Mathematics for Machine Learning, Probability and Statistics, Artificial Intelligence Programming, Computer Vision and Image Processing, Natural Language Processing
- **CS Theory**: Data Structures and Algorithms, Advanced Algorithms, Discrete Mathematics, Theory of Computation, Computational Complexity, Algorithmic Information Theory, Special Topics in Data Compression
- **Mathematics courses**: Linear Algebra, Real Analysis, Complex Analysis, Differential Equations, Mathematical Logic
- **Systems**: Operating Systems, Computer Networks, Principles of Programming Languages, Compiler Design, Database Management Systems, Programming Tools and Techniques.

Complete list of courses at ashudeepsingh.com/courses.html.

Teaching

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.

Teaching Assistant for **Machine Learning for Vision** **(Spring 2015)**
Faculty Instructor: Prof. Vinay P. Namboodiri

Skill Set

- **Programming Languages** – C, C++, Python, Java, C#, R
- **Platforms** – Linux, Windows
- **Web Development** – HTML, CSS, PHP, JavaScript
- **Other Tools** – Shell Scripting, Lex, Yacc, Matlab, SQL, Octave, LaTeX, Visual Studio, Eclipse, Git, Weka

Positions of Responsibility

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)**

- As the *Computer Room and Reading Room Secretary*, maintained the Hostel Website, Hostel Computer Center & implemented an online lending/borrowing system for books.
- Managed the hostel administration along with 9 other members of the Hall Executive Committee.

Secretary, Hospitality Cell, Techkriti 2011 **(2011)**

Extra-Curricular Activities

- 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 friends location on a Google Map, using Google Map API, mySQL databases and *facebook* APIs, during the Hack day by *Yahoo!* HackU 2012.
- Cleared the A-level Certificate of **National Cadet Corps (NCC)** from *1, CHD Naval Unit NCC, Chandigarh in 2007* with an A-grade.
- Placed 2^{nd} in **Madmen (Video Ad Making Competition)** in *Spectrum, IIT Kanpur Films and Media Festival, 2012.*