

SHIKHAR SHARMA

shikhar.sharma@microsoft.com

Research SDE II, Microsoft Research

<http://www.shikharsharma.com/>

RESEARCH INTERESTS

- Deep Learning
- Attention models
- Machine Learning
- Task-oriented dialogue
- Recurrent Neural Networks
- Computer Vision

RESEARCH STATEMENT

My long-term research interest lies in advancing the capabilities of current artificial intelligence systems. I am highly interested in neural network architectures augmented with memory and attention mechanisms.

My recent research deals with image-generation and task-oriented dialogue systems, which stand to gain significant advantage from using memory and attention similar to how we humans do. Prior to this, I have worked on visual attention models for action recognition and video description. I have broad interests in deep learning research in natural language processing and computer vision, which I believe are important research areas on the path to true artificial intelligence.

EDUCATION

University of Toronto

M.Sc. in Computer Science with the Deep Learning and Machine Learning group

Sep'14 - Feb'16

GPA: 3.83/4

- Thesis title: *Action Recognition and Video Description using Visual Attention*
- Thesis supervisor: *Prof. Ruslan Salakhutdinov*

Indian Institute of Technology (IIT), Kanpur

B.Tech. in Computer Science

Jul'10 - May'14

GPA: 9.3/10

- Thesis title: *Speech Recognition using Deep Belief Networks and Hidden Markov Models*
- Thesis supervisor: *Prof. Harish Karnick*

PUBLICATIONS

[†]equal contribution

Theses

Shikhar Sharma. 2016. "Action Recognition and Video Description using Visual Attention." Masters Thesis, University of Toronto

Journal papers

Ashesh Jain, Shikhar Sharma, Thorsten Joachims, Ashutosh Saxena. 2015. "Learning preferences for manipulation tasks from online coactive feedback." *The International Journal of Robotics Research (IJRR)*, 34, 1296-1313

Conference and arXiv papers

Shikhar Sharma, Dendi Suhubdy, Vincent Michalski, Samira Ebrahimi Kahou, Yoshua Bengio. 2018. "ChatPainter: Improving Text to Image Generation using Dialogue." [arXiv:1802.08216](https://arxiv.org/abs/1802.08216) [cs.CV]

Hannes Schulz[†], Jeremie Zumer[†], Layla El Asri, Shikhar Sharma. 2017. "A Frame Tracking Model for Memory-Enhanced Dialogue Systems." *Proceedings of the 2nd Workshop on Representation Learning for NLP (RepL4NLP)*

Layla El Asri[†], Hannes Schulz[†], Shikhar Sharma[†], Jeremie Zumer[†], Justin Harris, Emery Fine, Rahul Mehrotra, Kaheer Suleman. 2017. "Frames: A Corpus for Adding Memory to Goal-Oriented Dialogue Systems." *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue (SIGDIAL)*

Shikhar Sharma, Layla El Asri, Hannes Schulz, Jeremie Zumer. 2017. “Relevance of Unsupervised Metrics in Task-Oriented Dialogue for Evaluating Natural Language Generation.” [arXiv:1706.09799](https://arxiv.org/abs/1706.09799) [cs.CL]

Shikhar Sharma, Jing He, Kaheer Suleman, Hannes Schulz, Philip Bachman. 2017. “Natural Language Generation in Dialogue using Lexicalized and Delexicalized Data.” International Conference on Learning Representations (ICLR) Workshop

Shikhar Sharma, Ryan Kiros, Ruslan Salakhutdinov. 2016. “Action Recognition using Visual Attention.” International Conference on Learning Representations (ICLR) Workshop

Ashesh Jain, Shikhar Sharma, Ashutosh Saxena. 2016. “Beyond geometric path planning: Learning context-driven trajectory preferences via sub-optimal feedback.” Robotics Research: The 16th International Symposium (ISRR). Springer International Publishing, 319-338

Shikhar Sharma, Ryan Kiros, Ruslan Salakhutdinov. 2015. “Action Recognition using Visual Attention.” Neural Information Processing Systems (NIPS): Time Series Workshop

RESEARCH EXPERIENCE

Microsoft Research

Research SDE II

Feb'17 - Present

Montreal, Quebec, Canada

After Maluuba was acquired by Microsoft, I have done research on evaluation of natural language generation systems for task-oriented dialogue using unsupervised metrics. I have also worked on end-to-end trainable dialogue systems. I am currently working on systems that generate images based on text inputs (a caption or a dialogue). Apart from these, I have also worked on projects to integrate state-of-the-art research into Microsoft products.

Maluuba Research

Research Scientist

Mar'16 - Feb'17

Montreal, Quebec, Canada

I have worked on end-to-end trainable models for task-oriented dialogue systems with the dialogue group on Maluuba's Frames dataset. Prior to this, I worked on improving the natural language generation component of task-oriented dialogue systems over the then state-of-the-art by using lexicalized slot values aligned with delexicalized slots.

University of Toronto

Graduate Research Assistant

Sep'14 - Feb'16

Toronto, Ontario, Canada

Working together with Prof. Ruslan Salakhutdinov and Jamie Ryan Kiros, my masters thesis focused on soft-attention based recurrent neural networks for action recognition and video description generation.

Cornell University

Research Intern

May'13 - Jul'13

Ithaca, New York, USA

I worked with Prof. Ashutosh Saxena and Ashesh Jain on learning preferences over trajectories on robots such as the Baxter. Our approach required a non-expert user for training and the preferences we learned were governed by objects and human interactions in the environment.

AWARDS AND SCHOLASTIC ACHIEVEMENTS

- 2014** Department Entrance Scholarship (CS) of CAD\$10,000 at University of Toronto for 2014-2016
- 2013** Selected for the prestigious Cornell IIT-Internship Program, 2013
- 2012** IIT Kanpur Academic Excellence Award for 2011-12 for distinctive academic achievements
- 2011** O.P. Jindal Engg. and Mgmt. Scholarship for excellence in academics and leadership
- 2011** IIT Kanpur Academic Excellence Award for 2010-11 for distinctive academic achievements
- 2010** Awarded the CBSE Merit Scholarship for Professional Studies - AIEEE for 2010-2014
- 2010** All India Rank 434 in IIT-Joint Entrance Examination (amongst 0.48 million candidates)

KEY RESEARCH PROJECTS

ChatPainter: Improving Text to Image Generation using Dialogue *Nov'17 - Feb'18*
Dialogue group *Microsoft Research, Montreal, Canada*

- Implemented a stacked GAN model that generates images for a given caption and related dialogue
- Generation was done in 2 stages - the first generated a coarse 64x64 image; the second a refined 256x256 image
- Observed that adding an exploratory dialogue to a caption resulted in improvements over the baseline

Evaluation of Task-oriented Natural Language Generation systems *Feb'17 - May'17*
Dialogue group *Microsoft Research, Montreal, Canada*

- Implemented baselines and hierarchical recurrent models to generate natural language
- Set up a pipeline for evaluating generated sentences on word-overlap-based and word embeddings-based metrics
- Performed analysis of correlation between human scores and unsupervised metrics

Source code <https://github.com/Maluuba/nlg-eval/>

A dialogue system from human-human dialogues *Jun'16 - Feb'17*
Dialogue group *Maluuba Research, Canada*

- Organized and supervised Frames data collection and annotation with data analysts: 19,986 turns in total
- Generalized the state tracking task to frame tracking, which requires adding memory
- Designed and implemented tools to parse and preprocess the dataset
- Implemented a baseline and a natural language understanding module which extracts, anonymizes slot values

Project page <https://datasets.maluuba.com/Frames/>

A natural language generator for task-oriented dialogue systems *Mar'16 - May'16*
Dialogue group *Maluuba Research, Canada*

- Organized and supervised data collection using crowdsourcing and annotation with data analysts
- Used lexicalized slot values with slot labels in a recurrent encoder-decoder framework
- Trained a language model with similar decoder architecture and transferred weights
- Obtained improvement over state-of-the-art approaches on BLEU-4, METEOR, ROUGE_L & CIDEr scores

Action Recognition and Video Description using Visual Attention *Sep'15 - Feb'16*
Masters Thesis supervised by Prof. Ruslan Salakhutdinov *University of Toronto, Canada*

- Proposed a soft attention model to generate descriptions of videos in YouTube2Text dataset
- Used a Long Short-Term Memory based Recurrent encoder and decoder framework
- Built upon our previous work on action recognition published at NIPS and ICLR workshops

Mentions Talk by Prof. Gunhee Kim at NVIDIA GTCx, Seoul 2016: <https://goo.gl/LAJz9z>

Action Recognition using Visual Attention *Feb'15 - Aug'15*
Research Project with Prof. Ruslan Salakhutdinov *University of Toronto, Canada*

- Proposed a soft attention model to perform action recognition in video datasets like HMDB-51, Hollywood2
- Used GoogLeNet to extract features and an attention mechanism to dynamically pool the features
- Performed the classification task using a Recurrent Neural Network based on Long Short-Term Memory units
- Compared our models with relevant baselines and state-of-the-art Deep Learning based models
- Presented an extensive qualitative analysis of possibly how the attention model selected glimpses

Project page <http://shikharsharma.com/projects/action-recognition-attention/>

Source code <https://github.com/kracwarlock/action-recognition-visual-attention/>

Mentions Talk by Prof. Ruslan Salakhutdinov at Deep Learning Summer School, Montreal 2016

Speech Recognition using Deep Neural Networks and Hidden Markov Models *Jan'14 - Apr'14*
B.Tech. Project with Prof. Harish Karnick *IIT Kanpur, India*

- Used MFCC and its first and second order derivatives as features for the Deep Neural Network
- The DNN was integrated with a Hidden Markov Model which was used for the final classification
- Trained and evaluated on ISOLET and an IIT Kanpur dataset of roughly 120,000 Hindi audio recordings

Learning Trajectory Preferences for Manipulators via Iterative Improvement *May'13 - Jul'13*
Research Project with Prof. Ashutosh Saxena *Cornell University, USA*

- Learnt preferences over trajectories for 7 Degree of Freedom robot manipulators of Baxter Research Robot
- The algorithm requires the user to incrementally improve over the trajectory currently proposed by the robot
- Theoretical regret bounds of our algorithm match the asymptotic rates of optimal trajectory algorithms
- The robot was used to perform grocery checkout tasks

Project page <http://pr.cs.cornell.edu/coactive/>
YouTube demo <https://www.youtube.com/watch?v=uLktpkd7ojA>
Mentions Cornell University CS News (<https://goo.gl/owFPss>),
TechCrunch (<https://goo.gl/y2hW12>), KurzweilAI (<https://goo.gl/KyfN9U>)

Logic Studio (collaboration of Microsoft Research, Redmond and IIT Kanpur) *Apr'12 - Jul'12*
Research Project with Dr. Sumit Gulwani, Prof. Amey Karkare, Prof. Subhajit Roy *IIT Kanpur, India*

- Focused on creating an Intelligent Tutoring System for teaching First Order Logic
- Utilized the Stanford CoreNLP Parser to parse and retrieve meta-data, dependencies between words
- Generated logical formulae consisting of appropriate quantifiers, first order logic predicates and connectives
- Generated new questions by replacing determiners in sentences from our corpus

TECHNICAL SKILLS

Programming Languages Python, C, C++
Machine Learning tools PyTorch, Keras, Theano, Tensorflow, Matlab
Web Programming HTML, PHP, CSS, JavaScript, JQuery, MySQL

TEACHING EXPERIENCE

2015 Teaching Assistant for CSC411: Introduction to Machine Learning at University of Toronto
2015 Teaching Assistant for CSC321: Introduction to Neural Networks at University of Toronto
2014 Teaching Assistant for CSC309: Programming on the Web at University of Toronto

LINKS

Website <http://www.shikharsharma.com/>
GitHub <https://github.com/kracwarlock>
LinkedIn <https://ca.linkedin.com/in/sharmashikhar>
Scholar <https://scholar.google.ca/citations?user=bLA7DYwAAAAJ>

EXTRA-CURRICULAR ACHIEVEMENTS AND ACTIVITIES

- Secured third place among 275 participants in Firewall (ethical hacking competition) in Phoenix 2013 - Delhi Technological University's Technical Festival
- Awarded the Judges Choice Award at Yahoo! HackU 2012 for a "Tweet Map and Sentiment Analyzer" built on node.js, socket.io, Google APIs and JavaScript
- Secured second place in Coldfire (ethical hacking competition) in Effervescence 2012 - IIIT Allahabad's Technical cum Cultural Festival
- Winner, Infobahn (Web Designing competition) at Spectrum 2012, IIT Kanpur

- Held lectures on programming and organized programming competitions for Programming Club at IIT Kanpur.
- Organized and prepared the problem statements for CodeSprint, an ethical hacking competition which saw a lot of participation from IIT Kanpur students

VOLUNTEER EXPERIENCE

MariHacks

Feb'18

Mentor and Judge

Montreal, Quebec, Canada

MariHacks was a day-long hackathon for CEGEP and high school students where they had to create revolutionary software projects. I was one of the mentors who helped the learners. I was also one of the judges who had to evaluate the projects on Originality, Usefulness, Polish, and Technical Difficulty.

Ladies Learning Code

Sept'15

Mentor

Toronto, Ontario, Canada

Toronto: Data Insights with Python for Beginners was a hands-on, beginner-friendly python workshop. Beginners learnt basic programming fundamentals and also how to read and manipulate large data sets to draw insights for marketing and analysis. I was one of the mentors who helped the learners.

Lighthouse Labs

Feb'15

Mentor

Toronto, Ontario, Canada

The HTML500 was a one-day event where 50 of Canada's top tech companies came together to teach 500 people how to code, for free. I was one of the mentors who helped out beginners in learning HTML and CSS and creating a web page.