Sumeet Singh   **github.com/sumeet-iitg**

(412) 626-0636
sumeet.iitg@gmail.com

EDUCATION	**Carnegie Mellon University, Pittsburgh, PA**	
	M.S. in Language Technologies, School of Computer Science, Advisors: Prof. Jaime Carbonell, Prof. Alan W. Black	August 2019
	• GPA: 3.81/4.33	
	Indian Institute of Technology, Guwahati, India	
	B.Tech (Bachelors) in Computer Science and Engineering,	May 2013
	• Department Rank: 4/68, GPA: 9.36/10	

GRADUATE RESEARCH PROJECTS

Named Entity Recognition and Linking in Insurance domain.   *2018 – present*

- Implemented methods for efficient linking of 5 million people and 1 million vehicle records
- Designed heuristics for parsing important facts from unstructured and noisy insurance claim logs
- Working closely with an auto-insurance company to develop a comprehensive network fraud detection tool.

Evaluating Influence Functions for Memory Replay
in Continual Learning   *2019 – present*

- Implemented a Memory-replay technique to overcome catastrophic forgetting in neural networks.
- Experimented with novel "Influence-function" based sampling techniques for selecting points to keep in memory.

Improving Machine Translation systems on Noisy Text.   *2018 – 2019*

- Generated synthetic noise while training the model to simulate noise on social media.
- Improved performance on French-English by 5 BLEU points.

Language Modeling for Hindi-English Code Mixed Text .   *2017 – 2018*

- Implemented a multitask Neural language-model for code-mixed text with improved accuracy over state-of-the-art language model.
- Scraped and sanitized data from websites containing Hindi-English code-mixed text.

WORK EXPERIENCE

Software Engineer II, Microsoft, India   *2013 – 2017*

Developed features for Microsoft Azure's Disaster Recovery Service On-Premises Agent
- Replication of VM-Ware Virtual Machines to Azure.
- Over-the-wire compression during data transfer reducing the cost of data transfer by half.
- Enabled telemetry for quicker diagnosis and mitigation of customer issues.
- Conceptualized Azure Cloud Service for cross-geo replication of Azure VMs.

PUBLICATIONS

Singh, S., Vaibhav, Stewart, C. and Neubig, G. *Improving Robustness of Machine Translation through Synthetic Noise* NAACL 2019, *(link)*.

Singh, S., Chandu, K., Manzini, T. and Black, A. W. *Language informed modeling of code-switched text.* ACL Workshop on Code-Switching 2018, *(link)*.

Singh, S., Awekar, A. *Incremental shared nearest neighbor density-based clustering.* ACM CIKM 2013, *(link)*.

GRADUATE COURSEWORK

❏ Algorithms for NLP
❏ Neural Networks for NLP
❏ Neural Machine Translation
❏ Topics in Deep Learning

❏ Dialogue Systems
❏ Introduction to Machine Learning
❏ Probability and Mathematical Statistics

RELEVANT SKILLS

Programming: Python, Java, C/C++
Deep Learning: Pytorch
Virtualization: Hyper-V Virtual Machine, Virtual Hard Disk, Azure REST APIs
Version Control: GitHub, Team Foundation Server
Other tools: Pandas, Numpy, Scikit-Learn