Sanket Vaibhav Mehta

Ph.D. Student | School of Computer Science | Carnegie Mellon University
+1-412-626-1922 | svmehta@cs.cmu.edu | https://www.cs.cmu.edu/~svmehta | sanketvmehta

Research Interests

- Machine Learning, Natural Language Processing, Deep Learning, Optimization
- Focus areas: Learning from limited labeled data, multiple tasks, non-stationary data distributions (Continual/Lifelong Learning, Transfer Learning, Meta-Learning, Multi-Task Learning)

Education

Carnegie Mellon University, Pittsburgh — July 2020 - Present
Ph.D. student in Language Technologies, LTI, SCS — CGPA: **4.10/4.33**
Advisor: Prof. Emma Strubell

Carnegie Mellon University, Pittsburgh — August 2019 - February 2020
Ph.D. student in Language Technologies, LTI, SCS — CGPA: **4.08/4.33**
Advisors: Prof. Jaime Carbonell & Prof. Barnabás Póczos

Carnegie Mellon University, Pittsburgh — August 2017 - August 2019
Master of Science in Language Technologies, LTI, SCS — CGPA: **4.13/4.33**
Advisors: Prof. Jaime Carbonell & Prof. Barnabás Póczos

Indian Institute of Technology Roorkee, India — July 2011 - June 2015
Bachelor of Technology in Computer Science and Engineering — CGPA: **9.67/10.0**
President's Gold Medalist — Institute Rank: **1** (out of 987 students)

Research Experience

Carnegie Mellon University, Pittsburgh, PA — August 2017 - Present
Graduate Research Assistant

- My current work shows that generic pre-training alleviates the effects of catastrophic forgetting when learning multiple tasks sequentially compared to randomly initialized models, and it's due to finding flatter minima. Further optimizing for current task loss and loss basin sharpness to explicitly encourage flat loss basins during sequential fine-tuning leads to competitive results with the state-of-the-art methods.
- Previous work focuses on developing algorithms for simultaneous/sequential learning in neural networks from multiple tasks/ domains (EMNLP 2020), injecting inductive biases into deep models for inference-time domain adaptation (AAAI 2019) and semi-supervised learning with limited labeled data (EMNLP 2018).
- Developed machine learning models for part demand forecasting and part price prediction problems (as a part of Boeing/Carnegie Mellon Aerospace Data Analytics Lab academic research initiative).

Google AI, Pittsburgh, PA (Remote) — June 2021 - October 2021
Research Intern/ Student Researcher (Hosts: Jinfeng Rao, Yi Tay)

- We systematically study the problem of compositional generalization for data-to-text generation and propose a generic BLEURT based self-training approach for improving the model's generalization capabilities (ACL 2022).
- Contributed to a suite of 107 NLP tasks, where we show that massively multi-task pre-training can improve downstream performance on NLP tasks, overcoming trends of negative transfer between tasks while fine-tuning (ICLR 2022).

Adobe Research, Bangalore, India — June 2015 - July 2017
Member of Research Staff (Team Lead: Shriram Revankar, P. Anandan)

- Worked on designing algorithms for generating data-driven geo-fences to assist Adobe's digital marketing business and preventing inadvertent information disclosures by auto-tagging security policies.
- Transferred several technologies to Adobe Analytics (US Patent 9,838,843) and Adobe Experience Manager (US Patents 10,102,191 and 10,783,262).

Selected Publications

1. Improving Compositional Generalization with Self-Training for Data-to-Text Generation
   Sanket Vaibhav Mehta, Jinfeng Rao, Yi Tay, Mihir Kale, Ankur P. Parikh, Emma Strubell
   Annual Conference of the Association for Computational Linguistics (ACL 2022)
2. ExT5: Towards Extreme Multi-Task Scaling for Transfer Learning
   Vamsi Aribandi, Yi Tay, Tal Schuster, Jinfeng Rao, Huaixiu Steven Zheng, Sanket Vaibhav Mehta, Honglei Zhuang, Vinh Q. Tran, Dara Bahri, Jianmo Ni, Jai Gupta, Kai Hui, Sebastian Ruder, Donald Metzler
   International Conference on Learning Representations (ICLR 2022)
3. An Empirical Investigation of the Role of Pre-training in Lifelong Learning
   Sanket Vaibhav Mehta, Darshan Patil, Sarath Chandar, Emma Strubell
   Workshop on Theory and Foundation of Continual Learning (ICML 2021) (**Spotlight**)
4. Efficient Meta Lifelong-Learning with Limited Memory
   Sanket Vaibhav Mehta*, Zirui Wang*, Barnabás Póczos, Jaime Carbonell
   Conference on Empirical Methods in Natural Language Processing (EMNLP 2020)
5. Learning Rhyming Constraints using Structured Adversaries
   Harsh Jhamtani, Sanket Vaibhav Mehta, Jaime Carbonell, Taylor Berg-Kirkpatrick
   Conference on Empirical Methods in Natural Language Processing (EMNLP 2019)
6. Gradient-based Inference for Networks with Output Constraints
   Jay-Yoon Lee, Sanket Vaibhav Mehta, Michael Wick, Jean-Baptiste Tristan, Jaime Carbonell
   AAAI Conference on Artificial Intelligence (AAAI 2019)
7. Towards Semi-Supervised Learning for Deep Semantic Role Labeling
   Sanket Vaibhav Mehta*, Jay-Yoon Lee*, Jaime Carbonell
   Conference on Empirical Methods in Natural Language Processing (EMNLP 2018)
8. Preventing Inadvertent Information Disclosures via Automatic Security Policies
   Tanya Goyal, Sanket Vaibhav Mehta, Balaji Vasan Srinivasan
   Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2017)
9. An LSTM Based System for Prediction of Human Activities with Durations
   Kundan Krishna, Deepali Jain, Sanket Vaibhav Mehta, Sunav Choudhary
   The Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT 2017)

Issued Patents

1. Generating data-driven geo-fences (US 9,838,843)
2. Propagation of changes in master content to variant content (US 10,102,191)
3. Digital document update (US 10,489,498)
4. Tagging documents with security policies (US 10,783,262)
5. Digital document update using static and transient tags (US 10,846,466)
6. Tenant-side detection, classification, & mitigation of noisy-neighbor-induced performance degradation (US 11,086,646)

Relevant Courses

Introduction to Machine Learning (10-701), Algorithms for NLP (11-711), Neural Networks for NLP (11-747), Structured Prediction for Language (11-763), Deep Reinforcement Learning and Control (10-703), Multimodal Machine Learning (11-777), Deep Learning (10-707), Human Language for AI (11-724), Convex Optimization (10-725)

* denotes equal contribution

Accolades

- Department Research Fellowship, LTI, SCS (August 2017 - present).
- Recipient of **President's Gold Medal**, **Dr. A. N. Khosla Medal** and **Smt. Shashi Krishna Medal** for the Session 2014-15 for obtaining the highest CGPA amongst the B.Tech/ B.Arch/ IDD/ Int. M.Sc./ Int. M.Tech passing out students.
- Recipient of **Kathail Family Annual Excellence Award** and **Rakesh Agrawal Annual Excellence Award** presented by IIT Roorkee Heritage Foundation in 4^{th} year and 2^{nd} year of B.Tech in Computer Science and Engineering for outstanding Academic, Co-Curricular and Extra-Curricular achievements respectively.
- Recipient of **Certificate Of Trust Prize- Mr. Rai Singh Jain & Mrs. Shakuntla Devi Jain** presented by IIT Roorkee for the year 2014 and 2012 for The Student Obtaining Highest CGPA in B.Tech 3^{rd} year and 1^{st} year.

Responsibilities

- **Reviewer:** NeurIPS (2021, 2022), ICLR (2021, 2022), ACL ARR (2022), AAAI (2021), EMNLP (2019, 2020, 2021), ACL (2019, 2020), LTI SRS (2018)
- **Teaching Assistant:** (1) Algorithms for NLP (11-711), (2) Artificial Intelligence (15-681A/IITP-01) where responsibilities included conducting recitations, holding office hours, creating and grading assignments and exams.
- **Internship Mentor:** Mentored a group of 3 students at Adobe Research over the summer of 2016.
- **Founding Chair, IIT Roorkee ACM Student Chapter:** Presided over all the meetings of the chapter and of its Executive Council while serving as chair for IIT Roorkee ACM Student Chapter during the session 2014-15.
- **Ambassador for ACM:** Served as a campus ambassador for ACM at IIT Roorkee during the session 2014-15.