

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: **4.09/4.33**
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 (Theory of CL, ICML 2021).
- 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 Research, Mountain View, CA MAY 2022 - PRESENT
RESEARCH INTERN (HOSTS: [Yi Tay](#), [Jai Gupta](#))

- We introduce DSI++, a continual learning challenge for DSI that requires incrementally adding documents to the model and propose a two-fold approach focusing on training dynamics and data-based regularization to alleviate catastrophic forgetting.

Google AI, Pittsburgh, PA (Remote) JUNE 2021 - OCTOBER 2021
RESEARCH INTERN/ STUDENT RESEARCHER (HOSTS: [Yi Tay](#), [Jinfeng Rao](#))

- 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).

PREPRINTS/ UNDER REVIEW

1. An Empirical Investigation of the Role of Pre-training in Lifelong Learning
Sanket Vaibhav Mehta, Darshan Patil, Sarath Chandar, Emma Strubell
arXiv preprint arXiv:2112.09153
2. DSI++: Updating Transformer Memory with New Documents
Sanket Vaibhav Mehta, Jai Gupta, Yi Tay, Mostafa Dehghani, Vinh Q. Tran, Jinfeng Rao, Marc Najork, Emma Strubell, Donald Metzler
arXiv preprint arXiv:2212.09744
3. An Introduction to Lifelong Supervised Learning
Shagun Sodhani, Mojtaba Faramarzi, Sanket Vaibhav Mehta, Pranshu Malviya, Mohamed Abdelsalam, Janarthanan Janarthanan, Sarath Chandar
arXiv preprint arXiv:2207.04354

SELECTED PUBLICATIONS

1. Train Flat, Then Compress: Sharpness-Aware Minimization Learns More Compressible Models
Clara Na, Sanket Vaibhav Mehta, Emma Strubell
Findings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2022)
2. 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)
3. 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)
4. 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**)
5. 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)
6. 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)
7. 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)
8. 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)
9. 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)
10. 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)

^{*} denotes equal contribution

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)

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 4th year and 2nd 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 3rd year and 1st year.

RESPONSIBILITIES

- **Reviewer:** NeurIPS (2021, 2022), ICLR (2021, 2022, 2023), ACL ARR (2022), AAAI (2021), EMNLP (2019, 2020, 2021, 2022), ACL (2019, 2020, 2023), 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.