DEEP GANDHI

+1-(403)-401-2563 drgandhi@ualberta.ca deep1401.github.io deep1401 deep1401

EDUCATION

University of Alberta *2022 – 2024 (Expected)*
MSc. (Thesis) in Computing Science (Supervisor: Dr. Nidhi Hegde) Current GPA: **3.9/4**

Dwarkadas J. Sanghvi College of Engineering (University of Mumbai) *2018 – 2022*
Bachelor of Engineering (B.E.) in Computer Engineering CGPA: **9.55/10**

EXPERIENCE

University of Alberta Sep 2022 - April 2023
Graduate Teaching Assistant *Instructor: Dr. Nidhi Hegde*

- Worked as a Teaching Assistant of *Basics of Machine Learning* (Fall 2022) and *Ethics of Data Science* (Winter 2023).
- Responsible for conducting office hours, labs, grading assignments and exams, solving student queries, etc.

Unicode Research Aug 2020 - Dec 2022
Research Student *Advisor: Dr.Akash Srivastava, Swapneel Mehta*

- Worked on SimPPL to track online misinformation spread for adaptive governance (currently supported by NYC Media Lab, Wikimedia Foundation, and AI4ABM).
- Led a project for conducting audience analysis for Yale Daily News to track spread of articles on social media.
- Served as a TA for the Summer Machine Learning Course, UMLSC 2021, funded by **Google Research**.

JPMorgan Chase & Co. June 2021 - Aug 2021
Software Engineer Intern *Internship*

- Worked with the Investment Banking team to automate validation checks for every release using Python and pandas.
- Designed an automated system for evidence store creation of SNOW ticket files reducing the processing time from 90 mins to 10 mins.

Dwarkadas J. Sanghvi College of Engineering Jan 2021 - June 2021
Research Assistant *Advisor: Dr.Ramchandra Mangrulkar*

- Developed a project dealing with brain tumor segmentation using Federated Learning on the cloud.
- Trained UMLFiT & AWD-LSTM models for detection of Spear Phishing on a corpus of 73,000 emails.
- Worked on 2 chapters in a book by Chapman and Hall/CRC Press in the domain of Federated Learning & NLP.

Levyne Feb 2020 - May 2020
Machine Learning Engineer *Internship*

- Built the complete data analysis platform for the marketing team which performed RFM analysis on dynamic data.
- Responsible for building a chatbot using NLTK for customer interaction and a recommendation system using fast.ai.

PROJECTS

A Federated Approach for Hate Speech Detection *Guide: Dr. Zeerak Talat*

- Proposed a novel direction for detecting hate speech using Federated Learning with large landuage models.
- Observed improvements of around 10 F1 points using the proposed approach.

FineDeb: A Debiasing Framework for Language Models *Guide: Dr. Nidhi Hegde*

- Working on debiasing the training of various language models on isolated demographics such as race, gender, etc.
- Proposed a method which outperforms the existing approaches in terms of all stereotype metrics.

Automotive Component Failure Prediction *Guide: Dr. Kriti Srivastava*

- Collaborated with Deloitte to predict tyre life in vehicles using models such as MLP,XGB, etc.
- Designed a case study for the firm regarding tyre life uncertainty after extensive analysis of presented data.

Emoji Topic Prediction in Hindi Tweets using FL *Guide: Dr. Zeerak Talat*

- Cost sensitive learning and SMOTE for imbalanced emoji data using FedProx for training.
- Plan to release a dataset of around 200k tweets to predict emojis for resource constrained languages.
- Proposed a new algorithm for Federated Learning, 12x more efficient than some existing approaches.

NOTABLE PUBLICATIONS

[1] **Deep Gandhi***, Jash Mehta*, Nirali Parekh, Karan Waghela, Lynette D'Mello, and Zeerak Talat, "A Federated Approach to Predicting Emojis in Hindi Tweets," in *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing* ***(EMNLP)***. Association for Computational Linguistics, Dec. 2022, pp. 11 951–11 961. [Online]. Available: https://aclanthology.org/2022.emnlp-main.819

[2] Jay Gala*, **Deep Gandhi***, Jash Mehta, and Zeerak Talat, "A Federated Approach for Hate Speech Detection," in *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics* ***(EACL)***. Dubrovnik, Croatia: Association for Computational Linguistics, May 2023, pp. 3248–3259. [Online]. Available: https://aclanthology.org/2023.eacl-main.237

[3] Jash Mehta*, **Deep Gandhi***, Naitik Rathod, and Sudhir Bagul, "IndicFed: A Federated Approach for Sentiment Analysis in Indic Languages," in *Proceedings of the 18th International Conference on Natural Language Processing (ICON)*. NLP Association of India (NLPAI), Dec. 2021, pp. 487–492. [Online]. Available: https://aclanthology.org/2021.icon-main.59

[4] **Deep Gandhi**, Raghav Jain, Jay Gala, Jhagrut Lalwani, and Swapneel Mehta, "Expanding Access to ML Research through Student-led Collaboratives," in *Workshop on Broadening Research Collaborations in ML* ***(NeurIPS)***, 2022.

[5] Jash Mehta, **Deep Gandhi**, Govind Thakur, and Pratik Kanani, "Music Genre Classification using Transfer Learning on log-based MEL Spectrogram," in *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, 2021, pp. 1101–1107.

[6] **Deep Gandhi**, Jash Mehta, Nemil Shah, and Ramchandra Mangrulkar, "Federated Learning for Brain Tumor Segmentation on the Cloud," in *Cloud Computing Technologies for Smart Agriculture and Healthcare*. Chapman and Hall/CRC, 2021, pp. 261–278.

TECHNICAL STRENGTHS

Programming Languages:	Python, R, Javascript, C, C++
Libraries/Frameworks:	PyTorch, fast.ai, PySyft, Flower,Flask, FastAPI, numpy, pandas, scipy, Node.js
Databases:	SQL, MongoDB, Redis, Cloud Databases
Tools:	Git, Jupyter, Docker, Bash, Heroku, AWS, Azure, LaTeX

CO-CURRICULAR ACTIVITIES & ACHIEVEMENTS

1. Awarded the Alberta Graduate Excellence Scholarship 2022-23.
2. Contributed to the AI Career Accelerator Program at Amii to develop the AI201 course.
3. Worked on the Parrot platform demo with The Times (UK) at the JournalismAI Festival 2022.
4. Part of **Shalizi–Stats** reading group which focuses on the stats book "*Advanced Data Analysis from an Elementary Point of View*" by Prof. Cosma Shalizi and Bayesian Machine Learning.
5. Attended the Advanced Language Processing Winter School (ALPS) 2022.
6. Top 3 at *JPMorgan Chase Code for Good 2020* out of 75 teams.