

# DEEP GANDHI

📞 +1 (403)-401-2563 · ✉️ [drgandhi@ualberta.ca](mailto:drgandhi@ualberta.ca) · 🌐 [deep1401](https://deep1401.github.io) · 🌐 [deep1401.github.io](https://deep1401.github.io) · 🌐 [deep1401](https://deep1401.github.io)

## EDUCATION

### University of Alberta

MSc. (Thesis) in Computing Science, Supervisor: Dr. Nidhi Hegde

Aug 2022 - Present

GPA: **3.9/4.0**

## EXPERIENCE

### University of Alberta, Graduate Teaching Assistant

Sep 2022 - April 2023

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

### Unicode Research, Research student

Aug 2020 - Dec 2022

- Worked on [SimPPL](#) to monitor information 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 TA for **Google Research** funded 9-week Machine Learning Course UMLSC 2021 with 100+ students.

### Research Collaboration, Independent Researcher

Jan 2022 - Present

- Curated a dataset of ~1M tweets in low resource Hindi language & conducted emoji prediction using bi-LSTM, mBERT, IndicBERT, Hindi-Electra, XLM-R, etc. (**Accepted at EMNLP 2022**)
- Standardized 9 hate-speech datasets and experimented with FNet, DistilBERT, RoBERTa, etc. in federated & centralized settings. (**Accepted at EACL 2023**)
- Developed federated learning architecture for hate speech detection that obtained **14.52%** improvement in F1-score over traditional ML infrastructure, while minimizing risks to user-privacy.

### JPMorgan Chase & Co., Software Engineer Intern

June 2021 - Aug 2021

- Worked with the Investment Banking team to automate validation checks for weekly releases.
- 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, Research Assistant

Jan 2021 - June 2021

- Implemented 4 aggregation strategies for federated learning on non-iid medical data, using ResNet & U-Net.
- Trained UMLFiT & AWD-LSTM models for detection of Spear Phishing on a corpus of ~73k emails.

### Levyne, Machine Learning Intern

Feb 2020 - May 2020

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

## SELECTED PUBLICATIONS [*\* represents equal contribution*]

- [1] **Deep Gandhi\***, Jash Mehta\*, Niral 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, 2022 [📄](#).
- [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)*, Association for Computational Linguistics, 2023 [📄](#).
- [3] Jash Mehta\*, **Deep Gandhi\***, Naitik Rathod, and Sudhir Bagul, “IndicFed: A Federated Approach for Sentiment Analysis in Indic Languages,” in *Proceedings of 18th ICON 2021*, ACL Anthology, 2021 [📄](#).
- [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.

## PROJECTS

### 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 ICAT scores.

### Automotive Component Failure Prediction

Guide: *Dr. Kriti Srivastava*

- Collaborated with team at **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 data analysis.

## SKILLS

Languages:	Python, R, Javascript, C, C++
Libraries/Frameworks:	PyTorch, Tensorflow, Huggingface, Pandas, scikit-learn, NumPy, Flask, Node.js
Databases:	SQL, MongoDB, Redis, MySQL
Tools:	Git, Azure, GCP, Docker, Jupyter, Bash, Heroku, L <sup>A</sup> T <sub>E</sub> X