JASH MEHTA

🌐 jashmehta3300.github.io   ✉ jashmehta3300@gmail.com   in jashmehta3300

EDUCATION

Dwarkadas J. Sanghvi College of Engineering (University of Mumbai)   *2018 - 2022 (Expected)*
Bachelor of Engineering (B.E.) in Computer Engineering   Overall GPA: **9.60/10**
Applied Math, Machine Learning, Artificial Intelligence, Natural Language Processing, Algorithms & Data Structures

EXPERIENCE

Indian Institute of Technology (IIT), Gandhinagar   May 2021 - Present
Research Intern   *Advisor: Dr. Mayank Singh*

· Performed probing and robustness evaluation of automated review generators to adversarial perturbations.
· Formulated desiderata for an ideal review generator system and provided a public leaderboard along with a framework for unified & comprehensive measurement of their performance.
· Won **Best Presentation Award** at SRIP program & participated in discussions at the Computational Linguistics Group

Unicode Research   Aug 2020 - Present
Research Student   *Advisors: Dr. Akash Srivastava, Swapneel Mehta*

· **Projects:** Estimating the causal effect of mentorship on student career outcomes; and small-world simulation to model opinion polarization of online communities using Pyro.
· **Teaching Assistant:** Google Research funded 9-week Machine Learning Course UMLSC 2021 with 100+ students.
· Member of Shalizi-Stats reading group which deals with *Bayesian Statistics* and book by *Cosma Shalizi*.

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

· Made a project dealing with the brain tumor segmentation using Federated Learning on the cloud.
· Worked on a research project to identify Spear Phishing using low computational NLP approaches.
· Published 2 chapters in the domains of Federated Learning and Natural Language Processing.

Feople Org   July 2020 - Dec 2020
Machine Learning Intern   *Mumbai, India*

· Devised dynamic pricing strategy for a restaurant which helped in improving the sales by 28%.
· Worked with ML team for deployment of recommender system pipeline and other internal ML systems.

Margosatree Technologies   Jan 2020 - June 2020
Software Developer   *margosatree.com*

· Used Flask, MongoDB & Pandas to gain useful insights from large-scale dynamic data from manufacturing machines.
· Worked on a diverse array of client and internal projects like dynamic report generation using Selenium and Pandas.

RESEARCH & PUBLICATIONS

[1] **Jash Mehta***, Deep Gandhi*, Naitik Rathod, and Sudhir Bagul, "IndicFed: A Federated Approach for Sentiment Analysis in Indic Languages," in *Proceedings of ICON 2021: The 18th International Conference on Natural Language Processing*, ACL Anthology, *Accepted*.

[2] Deep Gandhi*, **Jash Mehta***, and Pranit Bari, "Ablation Analysis of Seq2Seq Models and Vanilla Transformers for Spanish to English Translation," in *Proceedings of the 3rd International Conference on Advances in Distributed Computing and Machine Learning*, Springer Nature, *Accepted*.

[3] **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)*, pp. 1101–1107, IEEE, 2021.

[4] 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*, pp. 261–278, Chapman and Hall/CRC, 2021.

[5] Deep Gandhi, **Jash Mehta**, and Ramchandra Mangrulkar, "Detection of Spear Phishing using Natural Language Processing," in *Cyber Security Threats and Challenges facing Human Life*, ch. 9, Chapman and Hall/CRC, *Accepted*.

PROJECTS

A Federated Approach to Predict Emojis in Hindi Tweets   *Guide: Dr. Zeerak Talat*

· Conducted cost sensitive learning and SMOTE for imbalanced emoji data using FedProx for training.
· Curated a dataset of ∼1M tweets in low resource Hindi language & conducted emoji prediction using bi-LSTM, mBERT, IndicBERT, Hindi-Electra, XLM-R, etc

Cross-Dataset Generalization for Hate Speech Detection using Federated Learning   *Guide: Dr. Zeerak Talat*

· An extension of Fortuna et al. to perform better cross-dataset generalization using Federated Learning.
· Experimented with AWD-LSTM, BERT, FNet, DistilBERT, etc. in federated & centralized settings.

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

· Collaborated with a Big4 Consultancy firm 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.

FedHealth   *Guide: Prof. Lynette D'Mello*

· Bachelor's thesis leveraging blockchain-based FL architecture for patient prognosis and hospital in-flow prediction.
· Created representations for personalized prescriptions based on user reviews using Med-BERT embeddings.

TECHNICAL STRENGTHS

Languages:	Python, R, Javascript, C, C++
Libraries/Frameworks:	PyTorch, Huggingface, Pyro, fast.ai, PySyft, Keras, Pandas, scikit-learn, NumPy
Databases:	SQL, MongoDB, Redis, PostgreSQL
Tools:	Git, Jupyter, Docker, Bash, Heroku, AWS, Azure, LaTeX

ACHIEVEMENTS

1. Awarded Inspire Scholarship, **Top 1%** candidates in the state for Higher Secondary Certificate (12th Grade), 2018.
2. **Coding member, DJS Arya:** a cansat team which achieved **24th rank** in PDR and **34th rank** (Worldwide) in 2018-19.
3. **Mentor at DJ Unicode:** jointly mentoring a team of ∼70 sophomores and juniors with selections for Google Summer of Code, international internships, and hackathon winners across the country.
4. **Top 8** at DJACM LoC 2020 Hackathon.
5. **Top 48** teams countrywide for "Project Deep Blue", an initiative by Mastek to solve real-world problems.
6. Presented various *paper reviews* as a part of the **Unicode Research Group**.
7. Selected for Advanced Language Processing Winter School (ALPS) 2022.

EXTRA-CURRICULAR ACTIVITIES

1. Avocations including hiking, long-distance running, chess, and guitar.
2. Captain of college table tennis team; won 5 tournaments in a year.