Rohit Mujumdar

rohit.mujumdar@yahoo.com | 470-530-9621| www.rohitmujumdar.com

Work Experience

Intel Corporation - *Research Engineer-Scientist* (Santa Clara, CA) *Nov 2022 – Present*

- **Data Science (Carbon-aware Computing)**: Bridged a critical gap in Intel's data-center carbon consumption project by identifying crucial factors affecting energy consumption. Built a Splunk→ MongoDB→ Pandas data collection pipeline. Built ML models (e.g. Regression, Random Forest) to predict power consumption from CPU/GPU workload features.
- **Data Science for Intel Simulator**: Built a comprehensive dataset by parsing & aggregating simulator results from diverse sources. Identified KPIs of Intel PC configurations using statistical methods (PCA, K-means, cosine-similarity)
- **Distributed ML Research**: Developed a novel multi-model, multi-node GNN training system using PyG & PyTorch DDP. Implemented a new strategy to share training batches amongst models to improve efficiency (to be open-sourced)
- **GenAI Application:** Implemented a local RAG-based solution for internal document interaction using Ollama models & Chroma Vector DB. Led POC from ideation to implementation, now exploring LLM evaluation frameworks.

NCR Corporation - *Software Engineer 1* (Atlanta, GA) *July 2021 - Nov 2022*

Developed & tested several REST API features for NCR's Ordering Platform using Java SpringBoot, PostgreSql, Google Cloud, Postman, Docker, Git. Presented features regularly to executive leadership, garnering team visibility.

IBM Research - *Research Intern* (Almaden Research Lab, CA) *June 2020 - Aug 2020*

- Engineered a time series model for predicting ticket resolution times, accuracy > 80% for 2.3+ million tickets.
- Led meetings with stakeholders to discuss and integrate feedback, ensuring continuous improvement and alignment of solution with business needs; Filed ***two patents*** with the USPTO showcasing potential market impact.

Georgia Institute of Technology – Graduate Researcher (Atlanta, GA) *Jan 2021 – May 2021*

- Investigated and identified fraud attack vulnerabilities in Twitter's anti-misinformation tool 'Community Notes'
- Proposed + built a more robust Note-ranking solution; research directly influenced/used in current Twitter algorithm**.**

Froot Research - *Data Scientist* (Pune, India) *Sept 2018 - May 2019*

- Developed a constantly evolving ML model for incident ticket clustering & trend forecasting for an operations analytics dashboard. Formulated metrics to evaluate cluster quality & ticket-cluster classification.
- Innovated a novel binning technique to enhance K-Means clustering, successfully delivering POC to client Atos.

HSBC Software Development India - *Software Engineer* (Pune, India) *Aug 2017 - Aug 2018*

- Built a Java application to interface with backend mainframe systems; reduced database query operations time by 50%

Technical Skills

Languages: Python (Pandas, NumPy, Scikit-learn, HuggingFace, NetworkX, SQLAlchemy, Flask, Matplotlib), Java, SQL
Frameworks/Tools : Jupyter, PyTorch, Pytorch Geometric, PyTorch Distributed, SLURM

Education

Georgia Institute of Technology, Atlanta, GA *Aug 2019 - May 2021*
Master of Science, Computer Science, Specialization: Machine Learning + Social Computing

Vishwakarma Institute of Technology, Pune, India *July 2013 - May 2017*
Bachelor of Technology, Computer Engineering

Publications and Patents

ICDMAI'22, *Recognizing Similar Relationships Within Ontology to Fine Tune Ontology* *Sept 2022*
US20220270019A1, *Ticket-Agent Matching and Agent Skillset Development* *Aug 2022*
ICWSM'22, *Overcoming Language Disparity in Online Content Classification with Multimodal Learning* *June 2022*
US20220164744A1, *Demand Forecasting of Service Requests Volume* *May 2022*
ASONAM'21, *HawkEye: A Robust Reputation System for Community-based Misinformation Detection* *Nov 2021*
INFORMS'20, *A Heuristic Approach To Compute Ticket Resolution Time* **(Best Poster - Honorable Mention)** *Nov 2020*

Key Projects

Do Scientific Ideas from more Prestigious Universities Spread Faster?

- Simulated epidemiological models on Microsoft Academic Graph to investigate the imbalance in spread of ideas
- Formulated novel measures for the subjective concepts like institutional prestige & idea quality

Can Machines Detect if you are a Jerk?

- Fine-tuned BERT to assess if we can replicate the sentiments and biases shared by Redditors on the r/AITA subreddit
- Classified reddit posts as morally ethical/unethical; achieved best accuracy of 64%

Conference Paper Acceptance Prediction

- Engineered novel custom features of research papers to assess factors deciding their 'acceptability' to conferences
- Implemented supervised learning models on AllenAI's PeerRead dataset; achieved best accuracy of 65%

Leadership and Volunteer Work

Reviewer, CVPR'24, NAACL'24, SWPC'24 (Intel) *2024*
Vice President, Public Relations, iNCRedible Toastmasters Club, NCR Corporation *July 2022 - March 2023*
Head Teaching Assistant, AI, Ethics & Society *Jan 2020 - May 2021*
Teaching Assistant, Knowledge-Based AI *Aug 2019 - Dec 2019*