

# MITHESH RAMACHANDRAN

2136965070 • mitheshr@usc.edu • linkedin.com/in/259mit • github.com/259mit

## ABOUT

**Data scientist** with expertise in **NLP (LLMs)**, **Geospatial AI**, Machine Learning and Databases. Skilled in SQL-driven reporting, dashboarding, and analytics. **Winner – McGill McMedHacks (1st in Votes, 2nd in Visualization) and Pratyay Case Study (1st of 247)**. Creator of **MAHA**, an open source package in PyPI.

## EDUCATION

**M.S., Computer Science, University of Southern California, CA** Aug 2023 - May 2025

3.65 GPA, Relevant coursework: Geospatial Info Management, Advanced Data Stores, Databases, ML, NLP

**PGD, Applied Statistics, Indian Statistical Institute, IND** Aug 2022 - Aug 2023

8 GPA, Relevant coursework: Advanced Probability, Continuous Probability, Advanced Data Analysis

**B.Tech, Data Science, Narsee Monjee Institute of Management Studies, IND** July 2018 - May 2022

3.65 GPA, Relevant coursework: Machine Learning, Deep Learning, Cloud (AWS), Statistical Methods

## TECHNICAL SKILLS

**Data Analysis + Visualization + ML:** Flask, **PowerBI**, **Tableau**, R, **Python**, **Pytorch**, NLP, Tensorflow, Java, Dask, **Hadoop**, OSMNX, **Excel**, **Dashboards**, LLMs, **Cloud Computing**, **AWS QuickSight**

**Systems and Databases:** **SQL**, **PostgreSQL**, **PostGIS**, **Kubernetes**, **Docker**, **Flask**, C, C++, **Shell**, **Bash**

## PROFESSIONAL EXPERIENCE

**NEXDIG Lab, USC, Los Angeles, CA / Remote: Graduate Researcher (Volunteer)** February 2025 – Present

- Contributing to research in the domains of **LLMs**, **knowledge bases**, and geospatial AI.
- Developing systems that enable efficient semantic querying and natural language interaction over complex datasets.

**USC Marshall School of Business, Los Angeles, CA: Graduate Assistant** June 2024 – Present

- Conducting data analysis for Handshake, **analyzing trends**, and generating detailed presentations and reports to support **strategic planning** and decision-making at the Marshall Career Center.
- Enhancing event marketing initiatives and **driving student engagement** for more than 15 fairs and events by leveraging **Semi-supervised learning** techniques to bucket events for improved comparative analysis

**NielsenIQ, Mumbai, Remote: Data Scientist, Consumer Insights (Intern)** November 2021 – May 2022

- Engineered a **scalable**, novel, and **automated ML/NLP pipeline**, reducing text data processing time from **8 hours to 15 minutes**, boosting operational efficiency by 97 percent.
- Led a **cross-functional transliteration** project across 10 Indian languages, collaborating closely with the Tech team to standardize multilingual scripts, reducing reliance on outsourcing and **cutting operational costs**.

**IIT, Remote: Researcher – NLP and Speech Recognition** February 2021 – January 2022

- Developed an Automatic Speech Recognition (ASR) Pipeline based on Indic languages such as Tamil and Assamese using ASR toolkits based on **Deep Neural Networks**, **Mixture Models**, and **Hidden Markov Models**.
- Led ASR research team to construct a **Text-To-Speech** Classifier in Kaldi achieving over **90 percent accuracy**.

## PROJECTS

**Spatiotemporal Traffic Forecasting and Spatial Indexing Optimization** Oct 2024 – Dec 2024

- Trained and optimized a **DCRNN** in PyTorch, improving Los Angeles traffic route forecasting accuracy by 10%.
- Designed an in-memory spatial network index, enhancing nearest-neighbor query performance by **2x** with IER.

**PostIndex: PostgreSQL Automatic Indexing Worker** Mar 2025 – May 2025

- Boosted throughput by up to **35x** by building a lightweight background worker for autonomous index tuning.
- Reduced query latency by up to **30x** on complex and nested workloads through cost-aware index management.

**Dynamic Gesture Feature Extraction for Word-Level Sign Language** Sep 2023 – Nov 2023

- Reduced dimensions by **75 percent** using Spectral Clustering on WLASL1000 and improved Pose-GRU results.