

OMKAR AURANGABADKAR

San Francisco, CA 94539

☎ +1 (669)-977-9156

✉ omkaraurangabadkar13@gmail.com

in LinkedIn

🐙 Github

📁 Portfolio

EXPERIENCE

Infosys Ltd (Client: T-Mobile)

Aug 2025– Jan 2026

Data Engineer

Kansas, USA

- Developed and maintained **15+ Power BI reports and dashboards** for T-Mobile stakeholders, transforming large-scale operational datasets into actionable insights, **reducing manual reporting effort by 30%** and improving decision turnaround time.
- Designed a **JIRA complexity scoring system** using **semantic text analysis and NLP techniques** to evaluate ticket scope and effort, **improving sprint planning accuracy by 20%** and enabling better workload prioritization across Agile teams.

Taral AI

Jun 2024 – Aug 2024

Machine Learning Engineer Intern

Houston, TX

- Engineered and maintained scalable **batch and real-time data pipelines on AWS SageMaker**, optimizing data ingestion and model training, which **reduced deployment time by 17%**.
- Designed a modular, prompt-driven SaaS system using **Python and LLMs** to automate unit test generation, **cutting manual workload by 24%** and improving developer velocity.
- Integrated and operationalized LLM-based models (**Claude AI, ChatGPT-4, Gemini 1.5 Pro**) in production environments with **95% uptime**, enhancing system stability and benchmarking rigor.
- Collaborated cross-functionally to iterate on system design, **applying formal testing, debugging, and monitoring techniques** to improve model performance and ensure data consistency.

Persistent Systems

Mar 2021 – Jul 2023

Associate Data Analyst

Pune, India

- Developed and maintained robust **SQL-based ETL pipelines** to cleanse and transform **large-scale financial datasets**, improving reporting accuracy by **14%**.
- Designed and owned **interactive Tableau dashboards** using **calculated fields, parameters, filters, and KPI views** to translate business requirements into insights, improving stakeholder visibility into financial performance.
- Leveraged **Airflow and Spark** to automate data workflows and reduce manual intervention, enabling **faster data delivery and decision support**.
- Performed KPI modeling and trend analysis directly in **Tableau**, enabling executive-ready visual reporting and **improving reporting accuracy by 22%**.
- Practiced Agile methodologies to rapidly deliver data products aligned with business needs, achieving a **90% satisfaction rate** through continuous feedback and iteration.

LinkCode Technologies

Mar 2019 – Oct 2020

Data Science Intern

Pune, India

- Designed data pipelines and conducted analysis using **Python, Pandas, and Scikit-learn**, improving segmentation accuracy by **19%** for targeted customer campaigns.
- Built **Tableau dashboards and visual reports** to analyze customer behavior and campaign performance, enabling data-driven decisions and improving **operational efficiency across teams**.

PROJECTS

AI-Powered Mock Interviewer | Prompt Engineering, LLMs, GenAI

Jan 2025 - Apr 2025

- Engineered a data ingestion pipeline to scrape and **normalize 10,000+ interview questions** from platforms like Glassdoor and GeeksforGeeks using **Groq API, Docker, and AWS Lambda**, enabling large-scale training of domain-specific LLMs.
- Built and deployed a voice-interactive AI interview platform using **AssemblyAI for real-time transcription** and **Google TTS for speech synthesis**, delivering real-time, personalized mock interviews via a NextJS frontend.

Wildfire Prediction in the USA | PowerBI, Data Mining

Jun 2024 - Dec 2024

- Demonstrated a **Power BI dashboard** that integrated data sources to **enhance real-time wildfire risk visualization, overcoming accessibility challenges and improving disaster response efficiency by 15%**.
- Enhanced the wildfire risk dashboard's interactivity to **overcome static data challenges, improving decision-making**. Added slicers, mapped high-risk zones, and identified key factors, **resulting in more vital insights and better risk assessment**.

EDUCATION

San Jose State University

Aug 2023 - May 2025

Master of Science in Applied Data Science

San Jose, CA

Coursework: Machine Learning, Big Data, Database Systems, Data Visualization, Deep Learning

University of Pune

May 2016 - Apr 2020

Bachelor of Engineering in Information Technology

Pune, India

Coursework: Machine Learning, Database Management, Cloud Computing, Computer Organization

SKILLS

Languages: Python, SQL
Data Engineering: Apache Spark (PySpark), Apache Airflow, AWS, Real-time Batch Pipelines
Data Management: Pandas, Numpy, MongoDB, MySQL, PostgreSQL
Analytics & BI: Tableau, Power BI, Matplotlib, Seaborn
Software Engineering: Git, Modular Code Design, Debugging, Testing
Machine Learning: Scikit-learn, PyTorch, LangChain, NLTK