

# OMKAR AURANGABADKAR

San Francisco, CA 94539

☎ +1 (669)-977-9156

✉ omkaraurangabadkar13@gmail.com

in LinkedIn

🐙 Github

## EXPERIENCE

### Taral AI

Jun 2024 – Aug 2024

Machine Learning Engineer Intern

Houston, TX

- Developed an LLM-powered SaaS toolchain to automate unit test generation, **reducing manual workload by 40%** and **showcasing LLM's reasoning capabilities through prompt-driven code comprehension**.
- Built and deployed scalable ML pipelines on Amazon SageMaker for data preprocessing and model training optimization, **achieving 43% faster deployment cycles and enhancing data quality**.
- Integrated and maintained AI models (Claude AI, ChatGPT-4, Gemini 1.5 Pro) in production with **95% uptime, delivering robust evaluation documentation for scalable inference and benchmarking**.
- Led performance benchmarking of Claude AI, ChatGPT-4, and Gemini 1.5 Pro through prompt refinement strategies to **enhance reasoning capabilities**.

### Persistent Systems

Mar 2021 – Jul 2023

Associate Data Analyst

Pune, India

- Executed complex SQL queries to clean, transform, and manage large-scale financial datasets, **improving data accuracy by 48%**.
- Crafted custom Tableau dashboards to track financial metrics and investment performance, **streamlining reporting workflows**.
- Collaborated with stakeholders to translate complex metrics into interactive dashboards using advanced SQL and ETL pipelines, **enhancing strategic decision-making**.
- Enabled strategic financial planning by integrating trend analysis and KPI monitoring, **resulting in a 52% boost in data accuracy**.
- Applied Agile methodologies to iterate on client requirements, **achieving a 90% customer satisfaction rate through adaptive financial solutions**.

### LinkCode Technologies

Mar 2019 – Oct 2020

Data Science Intern

Pune, India

- Used Python, Pandas, and machine learning algorithms to clean and analyze large datasets, **improving customer segmentation accuracy by 59%**.
- Applied business intelligence techniques to **enhance operational efficiency and data-driven decision-making**.

## 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

**Analysis & Visualization Tools:** Tableau, Power BI, Excel (VLOOKUP, Pivot Tables, VBA), Looker, Alteryx.

**Programming Languages:** R, Advanced SQL, Python, Java, Linux Shell, C#, C++.

**Database Management Systems:** MySQL, Oracle, MongoDB, CouchDB, AWS DynamoDB, AWS Redshift, Google BigQuery, NoSQL, Spark, Kafka, Kubernetes, Docker, Hadoop.

**Data Skills:** ETL Pipelining, Data Modeling, Data Presentation, Data Warehousing, Statistics, Apache Airflow.

**Data Science:** Predictive Modeling, A/B Testing, Causal Inference, Bayesian Methods.

**Cloud Technologies/Data Warehouse:** AWS, Google Cloud Platform (GCP), Azure, Snowflake.

**Tools/Services:** Git, REST API, RESTful Services, JSON, DevOps, Postman, Slack, PyCharm, JIRA.

**Other Skills:** Strong Analytical Skills, Problem-Solving, Collaboration, Business Communication.