ARCHIT MANEK

Bellevue, WA 98004 | architmanek@gmail.com | +1 617 961 2045 | github.com/archit-manek

WORK EXPERIENCE

Oracle Corporation   **Boston, MA**
Software Engineer   February 2023 – June 2025
- Developed and optimized APIs for managing study and tenant data, streamlining backend integration and simplifying access to clinical metadata across systems using Java and SQL.
- Boosted unit test coverage from ~35% to 80% across critical modules using JUnit and JMockit, significantly reducing regression bugs during release cycles.
- Designed and implemented database schema changes using Liquibase to ensure smooth, version-controlled updates to production databases, improving system integrity and minimizing downtime during releases.
- Automated audit record collection and comparison processes using SQL, reducing manual verification time by 40% and improving data integrity.
- Collaborated across teams in India, the U.S., and Europe to troubleshoot data inconsistencies and improve backend stability for global clinical trial platforms.

Oracle Corporation   **Boston, MA**
Software Engineer Co-op   January 2022 – June 2022
- Designed and built a global dashboard using Grafana and Prometheus to monitor key metrics such as active tenants and active studies for Oracle's Clinical One product, improving operational visibility.
- Developed UI components for customer-training webpage of OJET app to significantly improve user experience.

TECHNICAL SKILLS

Programming Languages: Python, Java
Data/ML: Polars, Pandas, NumPy, scikit-learn, SQL, Streamlit
Technologies: Medallion architecture, ETL, Oracle Database, Grafana, Liquibase, Parquet

EDUCATION

Northeastern University, **Khoury College of Computer Sciences**   **Boston, MA**
Master of Science in Computer Science   December 2022
Concentration: Software Engineering

Relevant Coursework: Building Scalable Distributed Systems, Foundations of Software Engineering, Database Management Systems, Object-Oriented Programming, Discrete & Data Structures, Algorithms, Data Management & Processing

PROJECTS

Football (Soccer) Analytics Pipeline | **[github.com/archit-manek/football-pipeline]**
*Python, Polars, Parquet | **In Progress***
- Designed and implemented a modular data pipeline for football analytics, transforming raw match, event, lineup, and 360° tracking data into structured, ML-ready datasets.
- Developed scalable ETL workflows in Python/Polars with automated logging, schema validation, and clear stage separation (Bronze/Silver/Gold), following Medallion architecture principles.
- Engineered advanced spatial and possession-based features to support deep learning models and modern football analysis.
- Managing the project end-to-end, from design and implementation to documentation and workflow automation.
- Currently developing and training custom expected goals (xG) and expected threat (xT) models using engineered features from the Gold layer, enabling advanced performance and tactical analysis from open-source football data.

ClinicalOneGPT – AI-Powered Chatbot for Internal Oracle Hackathon
Python, Google Gemini API, RAG architecture, LLMs, Vector Databases, Streamlit
- Led the design and development of a Retrieval-Augmented Generation (RAG) chatbot to help users query Oracle ClinicalOne documentation via natural language.
- Built a pipeline for semantic search using document chunking, embeddings, and a vector database to enable accurate, real-time responses.
- Integrated Google Gemini LLM for proof of concept; proposed transition to OCI GenAI stack (embedding models, Cohere/Llama).
- Focused implementation on a complex section "Sponsors and CROs" spanning 900+ PDF pages, showcasing the chatbot's ability to simplify dense technical content.
- Collaborated in a 4-person team to deliver a working MVP under time constraints, using Streamlit to demo the chatbot interface and present the solution to Oracle stakeholders.