

Mahavir Chaudhari

mveer.rc@gmail.com | My Portfolio | LinkedIn | Github | +91 9405476969

Education

BITS Pilani, BE in Computer Science 2019-2023

- Coursework: Computer Architecture, Machine learning, Deep Learning, Database management, Operating Systems, Design and analysis of Algorithms, Data structure and algorithms.

Experience

Data Engineer II, Providence Global Center - Hyderabad 2023 Jul - Present

- Redesigned the Mandatory Days Off scheduling system using Snowflake SQL, Oracle DB, and Control-M, re-architecting allocation logic with optimized set-based queries and scheduling algorithms; **reduced run time by 92% (2 h to 10 m)** and improved daily workload balance for **6,000+ doctors and nurses**.
- Engineered a project-to-GL reconciliation pipeline in Snowflake, Control-M, and Azure Data Factory, consolidating 8 manual reports into a unified table that processes **1B+ financial records daily (\$400M+)**; **reduced engineer effort by 80%** and enabled 4+ teams to deliver audit-ready reports.
- Mentored an intern while architecting a Python + Streamlit + Snowflake test automation framework, **accelerating QA cycles by 12x (14h to 1h)** and enabling faster release pipelines.

Software Engineer Intern, Fortanix - Bangalore 2023 Jan - June

- Developed comprehensive unit tests in Rust for encryption and key management modules, **expanding code coverage by over 35%** and improving the reliability of confidential computing workflows.
- Integrated Rust test suites into CI/CD (GitHub Actions, Docker), reducing build time by 20%.

Machine Learning Engineer Intern + Part-time Work Streamn.ai, Inc. - Remote 2021 May - 2022 Aug

- Engineered 2 RESTful APIs (FastAPI, Google Cloud, AWS) **handling 10K+ weekly requests** supporting MLops.
- Optimized text recognition pipeline (Python, Bash), reducing execution time by 30% (5s to 3.5s) and developed a multimedia web service (MySQL, SQLite, Rockset, Flask, JavaScript) to generate image output for the incoming video stream, improving the speed of rollout by 25%.

Projects

Smart Research Scout

- Built a full-stack AI-driven **RAG platform using FastAPI, PostgreSQL, and Redis**
- Deployed with Docker and GitHub Actions; integrated React analytics dashboard (Plotly/D3).

Small Data (Data Summarizer)

- Built an automated data summarization and insight extraction tool using **Python, FastAPI, and LangChain, integrating OpenAI APIs, Pandas, and SQLAlchemy** for data processing, deployed with Docker.

Portfolio Website

- Built an interactive portfolio using JavaScript with Canvas APIs, for smooth, scroll-triggered visuals. Designed a responsive, mobile-optimized layout, with reduced-motion controls and dynamic style injection.

Technical Skills

- **Languages:** Python, SQL, Golang, Java, Rust, Bash, C/C++, JavaScript, HTML
- **Frameworks & Tools:** React, FastAPI, Flask, Streamlit, PySpark, dbt, PyTorch, NLP/OCR libraries
- **Data & Cloud:** Snowflake, Databricks, Delta Lake, Fivetran, Oracle DB, Azure Data Factory, AWS, GCP (BigQuery, Firebase), MySQL, PostgreSQL, SQLite, Docker, Airflow, Control-M
- **Data Engineering & Operations:** ETL/ELT Pipeline Design, LLMs, SLMs, RAGs, MCPs, Data Quality Automation, CI/CD (GitHub Actions, Jenkins), Kubernetes, Scheduling Systems
- **Analytics & BI:** Power BI, Data Visualization, Stakeholder Reporting