Nithya Chandra Mohan

nithya.chandramohan1@gmail.com | 646-425-1963 | Los Angeles, CA (Open to Relocate) | GitHub | LinkedIn

Education

California State University, Long Beach — Expected Graduation Date: Jan. 2027
M.S. in Computer Science
Relevant Courses: Database Management Systems (DBMS), Distributed Computing, Advanced Software Engineering.

Experience

Outamation (Extern Program) — Jan. 2026 – Present
Data Engineering Extern — Remote

- Builds **modular data ingestion and transformation pipelines** using **Python**, **OCR** (**Tesseract**, **PaddleOCR**), and **PDF parsing**, processing **200+ page unstructured documents**.
- Develops **document retrieval workflows** using **Python** and **LlamaIndex**, applying **chunking** and **metadata-based filtering** to support **semantic search** use cases.
- Performs **data quality checks** and **performance evaluation** on **OCR** and **retrieval pipelines**, documenting **findings** and **optimization considerations**.

Deloitte Australia (Forage Virtual Experience) — Jan. 2024 – Feb. 2024
Data Analytics Virtual Intern — Remote

- Analyzed **machine telemetry datasets** from **5+ global factory locations** using **Excel** and **Tableau**, identifying downtime and failure patterns that improved insight accuracy by **25%**.
- Built **interactive Tableau dashboards** with operational KPIs, enabling factory-level performance monitoring and reducing issue identification time by **30%**.

Aaishwarya Construction Company — Apr. 2022 – Nov. 2024
Data Operations Associate — Chennai, India

- Built and maintained **structured data ingestion and transformation workflows** using **Excel** and **Google Sheets**, improving **data accuracy** and **data retrieval efficiency**.
- Designed **ETL-style reporting pipelines** to process **labor, material, and cost data** using **Excel automation**, reducing **manual processing time** by **40%**.
- Performed **data quality checks**, **validation**, and **aggregation** on **multi-year operational datasets** to support **analytics** and **operational decision-making**.

Publications

The Role of Generative AI in Transforming Healthcare Services
IEEE – 25th International Conference on Software Quality, Reliability, and Security Companion (QRS-C), 2025
DOI: 10.1109/QRS-C65679.2025.00015

- Analyzed applications of **Generative AI in healthcare**, including diagnostics, clinical documentation, drug discovery, and **privacy and bias considerations**.

Projects

End-to-End Document Data Pipeline (Independent Project) | Python, Pandas, SQL, ETL, OCR — Jan. 2026 – Ongoing

- Built **data ingestion and transformation pipelines** in **Python** to process **structured and semi-structured data**, improving data readiness for analytics.
- Implemented **ETL workflows** with **data validation**, **schema enforcement**, and transformations using **Pandas and SQL**, reducing data inconsistencies by **30%**.
- Documented **data pipelines, schemas, and processing logic**, improving maintainability and reducing onboarding time by **25%**.

Automated Business Analytics & KPI Pipeline (Independent Project) | Python, SQL, Analytics — Oct. 2023

- Developed **ETL-style analytics workflows** using **Python** and **SQL** to process **structured datasets** for reporting and analysis.
- Performed **data quality checks** and **aggregation** to support accurate **KPI calculation** and validation.
- Optimized **data processing logic** to improve **query performance** and **reporting efficiency**.

Skills

Languages: Python, SQL
Data Engineering: ETL/ELT, Data Pipelines, Data Ingestion, Data Quality & Validation
Analytics & Libraries: Pandas, KPI Analysis, Data Aggregation, Query Optimization, Documentation, Git