

VATSA ARVIND KALA

College Station, Texas 77840

☎ (979) 739-8190 ✉ vatsakala.contact@gmail.com [in Vatsa kala](#) [🔗 Vatsakala](#) [📁 Portfolio](#)

EDUCATION

Texas A&M University, College Station, TX

Master of Science in Management Information Systems — GPA: 3.83/4.0

Aug 2024 - May 2026

College Station, Texas

Pandit Deendayal Energy University

B.Tech in Electronics And Communication Engineering — GPA: 9.42/10

Oct 2020 - May 2024

Gandhinagar, India

TECHNICAL SKILLS

Languages & Frameworks: Python, SQL, C++, R, Javascript, CSS, Embedded C, Flask, OOP, MATLAB

Tools & Tech Stack: AWS(EC2, Fargate, S3, RDS), Tableau, Power BI, GIT, JIRA, Postman, Kubernetes, Confluence, MERN

Databases & Interests: Snowflake, MySQL, PostgreSQL, MongoDB, SPSS, STATA, Linux, Machine Learning, Data Structures

Certifications: Professional Scrum Master 1, Fundamentals of Deep Learning, Microsoft Azure Data Engineer Associate

PROFESSIONAL EXPERIENCE

Trading Technologies

May 2025 - Dec 2025

Data Engineering Intern

New York, USA

- Collaborated on migration of **4 Petabytes** of trading data from MSSQL to PostgreSQL and Snowflake on AWS, designing **500+ table schemas** to boost performance and maintain data integrity during platform modernization
- Engineered backend APIs using **Flask and SQLAlchemy**, deployed on **AWS** with unit tests and **ELK-standard logging**, facilitating real-time analytics and supporting rollout of firm's next-generation TradeZoom platform

Texas A&M University

Jan 2025 - May 2025

Student Assistant

College Station, USA

- Elevated biomedical research by enhancing GNN and LLM models through advanced optimization techniques, attaining a **20% improvement** in model performance and a **15% increase** in predictive accuracy for critical applications
- Restructured data preprocessing and feature engineering workflows, reducing processing time by **30%** and increasing model reliability for high-impact biomedical applications

ARK Meditech Systems

Dec 2023 - Jun 2024

Data Intern

Ahmedabad, India

- Streamlined data pipelines with **Spark SQL** on Hadoop, processing **10,000+** records daily and improving efficiency by **18%**
- Developed **SQL-driven Tableau dashboards**, strengthening visibility into trends, compliance, and boosting insights by **30%**

E-infochips

May 2023 - Jun 2023

Research Intern

Ahmedabad, India

- Conducted research on **10+** smartphone systems, suggesting hardware optimizations advancing design by **15%**
- Led **5** engineers to implement predictive research, driving semiconductor design enhancements

PROJECTS

Point of sales database system 🔗 | Amazon Web Services, MongoDB, ETL, Cloud Architecture

- Designed a **high-performance ETL pipeline**, integrating **10,000+ data points** from CSVs into a relational DB, promoting query speed by **25%** through strategic optimization of views, schemas, triggers, and stored procedures on AWS EC2
- Transformed POS database scalability and performance using **MongoDB**, slashing latency and catalyzing a **50%** increase in transaction volume while maintaining system efficiency

Mai Shan Yun Analytics Dashboard 🔗 | SQL, Streamlit, Pandas, Business Intelligence

- Architected a full-stack analytics ecosystem that unified **13+ operational datasets** including POS transactions and inventory logs into an automated ETL workflow powering real-time KPIs, interactive dashboards, and demand forecasting
- Formulated an **AI-driven insight engine** using Claude integration that generated contextual business recommendations and increasing forecast precision by 27%. **Received Top-3 honors at TAMU Datathon 2025**

RegDoc: AI-Powered Regulatory Document Classifier 🔗 | Databricks, OCR, LLaMA 3.1, Data Governance

- Constructed a **multi-stage data classification pipeline** combining OCR extraction, heuristic safety checks, and dual LLM reasoning to label regulatory documents by sensitivity with 93% accuracy
- Synthesized **confidence-based calibration**, citation tracing, and **human-in-the-loop feedback** to deliver explainable, audit-ready compliance results while cutting manual review effort by 40%

LEADERSHIP AND EXPERIENCE

Student Member, Soft Computing Research Society

May 2023 - May 2024

Organized **4+ expert talks** on applications of CNNs in computer science

Ahmedabad, India

Club Member, Rotaract

Aug 2022 - May 2023

Spearheaded community welfare events such as **4 blood donation camps** and **6 food drives**

Ahmedabad, India