

# Akshat Baranwal

✉ [abaranw2@asu.edu](mailto:abaranw2@asu.edu) | 📞 +1-6028049027 | [linkedin.com/in/akshatbaranwal21/](https://www.linkedin.com/in/akshatbaranwal21/) | [Github.com/11mosfets](https://github.com/11mosfets)

## Work Experience

**Larsen & Toubro Infotech, India**

01/24 to 08/24

### Data Engineer

- Engineered an automated **data pipeline using Python scripting** to parse, shred, and load massive, semi-structured XML files (200,000+ records) into structured Snowflake tables.
- Optimized data ingestion and decreased processing time by 90%, enabling faster data availability for analysis.
- Developed and deployed internal **data visualization** dashboards to monitor pipeline health, track data quality metrics, and provide stakeholders with real-time process insights.

**Xagrot LLC, Akron, Ohio - Remote**

06/25 to 08/25

### Automation Engineer - Intern

- Solved a critical bottleneck in PCB validation by developing a **Python automation** suite that integrates KiCad and Salome, cutting the time required for thermal simulation setup by over **99 %**.
- Automated the entire pre-processing chain, from Gerber file import and 3D model generation to the final meshed geometry setup for FEA in ElmerFEM.

## Education

**Arizona State University, Tempe**

08/24 to 05/26

M.S. in Computer Engineering

**CGPA: 3.4**

Relevant Coursework: Fundamentals of Algorithm, Machine Learning with FPGA deployment, Semantic Web Mining

**RGPV University, India**

08/19 to 06/23

B.Tech. in Electrical Engineering

**GPA: 7.8/10 First Division With Distinction**

Relevant Coursework: Microprocessor and Digital Electronics, Discrete Mathematics, Data warehousing.

## Project Work

### • Cloud-Native Semiconductor Digital Twin & ETL Pipeline **Dashboard** (Summer 2025):

Architected a fully automated data pipeline on Oracle Cloud Infrastructure **OCI** to simulate 5nm fabrication process drift and environmental excursions. Engineered a **ETL** system running via Cron in a virtual environment to ingest continuous telemetry into an **MySQL** data warehouse. Developed a **real-time Streamlit analytics** suite featuring interactive SPC control charts and automated yield recovery tracking, achieving 24/7 observability without manual intervention.

### • Hybrid RAG E-commerce Recommendation System (Fall 2025):

Designed and implemented a standalone graph retrieval agent for a laptop product dataset, utilizing **Neo4j** to model complex relationships. Developed an LLM driven interface that functions as a reasoning layer on top of the graph database, enabling the system to autonomously translate natural language queries from a central **orchestrator** into optimized **Cypher queries**. Enhanced recommendation explainability by leveraging **graph traversals** to surface relational insights that structured SQL and vector searches could not capture alone. Collaborated on a **hybrid retrieval architecture**, creating a specialized agent that interoperates with SQL and Vector search components to provide comprehensive answers for e-commerce queries.

### • Smart Grid Energy Monitoring & Forecasting System (2025):

Developed a real-time IoT dashboard using Streamlit and Python to monitor PV array health, enabling instant visualization of voltage sags and thermal anomalies across distributed sensors. Architected a time-series data pipeline utilizing SQL to aggregate high-frequency sensor data, optimizing query performance for historical power generation analysis. Engineered a Grid Load Forecasting model using Random Forest to predict next-day energy demand, performed feature engineering with lag features and rolling windows to reduce Mean Absolute Error (MAE) by **12%**.

## Skills

**Languages:** Python, SQL, Cypher (Neo4j), Shell Scripting.

**Data Engineering:** ETL/ELT Pipelines, Snowflake, MySQL, Data Warehousing, XML Parsing, Web Scraping, Data Ingestion, Automated Pipelines.

**AI & Machine Learning:** Large Language Models (LLMs), Llama, Retrieval-Augmented Generation (RAG), Vector Search, Graph Retrieval Agents, Semantic Analysis, Machine Learning (CNN).

**Databases:** NoSQL (Neo4j), Relational Databases (SQL), Vector Databases.

**Cloud & Dev Tools:** Cloud Computing (OCI), Streamlit, Tableau, Cron/Linux Automation, Git/GitHub, CI/CD, Object-Oriented Design, Data Structures & Algorithms, Distributed Systems, Software Development Lifecycle (SDLC)