Dinesh Arumugam

darumugam001@gmail.com | LinkedIn | PortFolio | +1 (607) 297-8068 | Binghamton, NY

Professional Summary

Computer Engineering with experience in Data analytics, Machine learning, and deployment-ready pipelines. Built solutions using Python, SQL, Tableau, Power BI, TensorFlow, XGBoost, Docker, MLflow, and Kubernetes to improve accuracy, reduce latency, and automate data workflows with focus on responsible AI deployment and data privacy compliance.

Technical Skills

Languages: Python, SQL (PostgreSQL, MySQL), R, Rust
Data & Analytics: Pandas, NumPy, Excel, ETL, Data Cleaning, Feature Engineering, Time Series, Hypothesis Testing, A/B Testing
ML & AI: scikit-learn, TensorFlow, PyTorch, XGBoost, Anomaly Detection, RAG (LangChain), MLflow
BI Tools: Tableau, Power BI, Streamlit
Cloud/MLOps: AWS, Snowflake, Spark/PySpark, Hadoop, FastAPI, Docker, Kubernetes, CI/CD

Experience

Research Assistant **Dec 2024 – Present**
Binghamton University (SUNY), Binghamton, NY

- Built an EMG gesture-recognition pipeline for real-time use; packaged the inference flow as a lightweight module and aligned inputs/outputs with the hardware/firmware setup, cutting latency by 50%.
- Improved signal quality with preprocessing and feature work, and added simple runtime checks (logging + confidence thresholds) to catch noisy sessions; debugged failure cases and iterated on the model.

Data Scientist Intern **May 2024 – Aug 2024**
Universal Instruments Corporation, New York

- Built a Python computer-vision pipeline to detect semiconductor defects; worked with QA/process engineers to confirm labels and acceptance criteria, improving detection accuracy by 30%.
- Connected model outputs to a SQL data store and standardized the scoring workflow; used queries to spot low-confidence runs and recurring error patterns for follow-up.
- Reduced inspection time by 50% by optimizing preprocessing for raw sensor/image data and reproducing edge cases from real samples to fix issues quickly.

Data Analyst– remote **May 2022 – Aug 2023**
Emglitz Technologies, Coimbatore, India

- Built ETL workflows to process 100,000+ records, improving accuracy and reliability of datasets used for reporting. Tableau/Excel KPI dashboards and wrote short, plain-English summaries for business teams, improving decision metrics by 40

Projects

FraudGuard: Real-time Card Fraud Detection & Investigator Copilot

- Built a low-latency fraud scoring system (XGBoost + anomaly signals) with sub-100ms scoring; containerized with Docker and tracked experiments with MLflow (ROC-AUC > 0.95).
- Implemented a RAG + Qdrant workflow to retrieve fraud playbooks and reduce manual investigation time by 30%.

CareIQ: Readmission Risk Prediction & Clinical Copilot

- Built a 30-day readmission risk pipeline and served it via FastAPI; aligned the API response format for integration and deployed on Kubernetes (sub-500ms inference).
- Built a RAG + vector database workflow to synthesize EHR context and speed clinician interpretation by 70%.

Education

Binghamton University, SUNY **Aug 2023 – May 2025**
B.S. Computer Engineering
Relevant Coursework: Statistical Analysis, Predictive Analytics, Deep Learning, Reinforcement Learning

PSG College of Technology **Aug 2021 – May 2023**
B.Tech Computer Science and Engineering
Achievement: 30% scholarship for academic performance (1st & 2nd year)

Leadership & Certifications

Peer Mentor, Watson College Mentoring Program (Binghamton University) **Sep 2024 – April 2025**
Mentored 15+ first-year engineering students in programming, data science, and career development.
Certifications: IBM Deep Learning & Reinforcement Learning (Coursera); Complete Data Science Bootcamp (Udemy); Advanced ML (Binghamton University)