Erick Saul Means

1-(317)-478-0066 | saul.amepm@gmail.com | linkedin.com/in/saulm24 | github.com/itsSaul24
US Citizen | Bilingual - English/Spanish | Manhattan, NY

Education

Bachelor of Science in Data Science, Purdue University West Lafayette — *Graduated 2024*
Master of Science in Computer Science, Georgia Institute of Technology — *In-Progress*

Professional Experience

Data Science Analyst, Bank of New York, Manhattan, New York, NY — *August 2024 – Present*

- Lead end-to-end machine learning (ML) and data engineering projects for production services department, including production AI agent deployment, automated data processing systems, and managing stakeholder relationships.
- Used Python and Git CI/CD to develop and deploy production reasoning-based AI agent leveraging LLMs with advanced prompt engineering (COSTAR framework) and RESTful API integration, enabling automated insights extraction from complex enterprise data and reducing manual analysis workflows.
- Managed cross-functional stakeholder relationships for ServiceNow data requirements, designing custom ETL solutions and data ingestion points tailored to business needs.
- Applied data modeling principles using Oracle PL/SQL and Pentaho (ETL platform) to build automated data retention algorithms processing millions of daily records across multiple databases, ensuring data compliance and maintaining referential integrity at enterprise scale.
- Engineered comprehensive ETL pipeline using Oracle PL/SQL to create the enterprise ”golden source” for 50+ ServiceNow KPIs, establishing centralized data warehouse architecture that powers monthly governance dashboards used by stakeholders across the organization.
- Presented technical findings to executive leadership in monthly governance meetings, translating complex data insights into actionable business recommendations.
- Led intern program as project manager for Service Management department, guiding 4+ intern through full SDLC while assisting in architecting production AI agents.

Contract Data Scientist, Accenture, Remote — *November 2023 – July 2024*

- Automated SAP Signavio KPI analysis pipeline using Python-based ML solutions with NLP and RAG techniques, transforming meaningless raw KPI data into actionable business insights and eliminating manual categorization workflows for enterprise stakeholders.
- Engineered custom ranking system by screen-scraping SAP Signavio KPI data, applying transformer models (BART, Google T5) for feature engineering, and implementing RAG-based model with vector database storage (langchain) and OpenAI integration for intelligent KPI summarization.
- Used Python and Streamlit to create web application with end-to-end automation, integrating screen-scraped data with machine learning models for business stakeholders.

Data Science Intern, Bank of New York, Manhattan, New York, NY — *June 2023 – August 2023*

- Used Python to clean and transform dirty data for Oracle data warehouse migration, successfully preparing datasets for enterprise-level data integration and storage.
- Applied Tableau visualization expertise to build multi-layer Sankey chart demonstrating ServiceNOW parameter data flow, enabling stakeholders to understand complex data relationships and system interactions.

Projects

FinGraph: Financial Intelligence Platform | *PyTorch Geometric, GCP, MLOps* — *Spring 2025 – Present*

- Developing end-to-end financial intelligence platform that uses Graph Neural Networks to discover hidden relationships between financial entities and detect patterns invisible to traditional analysis methods.
- Applied PyTorch Geometric and Graph Neural Networks to model complex financial entity relationships, enabling advanced pattern recognition and anomaly detection in financial networks.
- Implemented Docker containerization and GCP Cloud Run to deploy serverless microservices architecture with CI/CD pipelines, ensuring scalable and production-ready MLOps practices.

March Madness Tournament Prediction | *Python, scikit-learn* — *Spring 2024*

- Built machine learning system to predict NCAA tournament outcomes using ensemble methods and feature engineering, successfully forecasting basketball tournament performance.
- Applied Python with scikit-learn algorithms including Logistic Regression, SVM, KNN, and Random Forest to analyze team performance data, achieving a top 67% accuracy in predicting tournament results.