

Erick Saul Means

1-(317)-478-0066 | saul.amepm@gmail.com | [linkedin.com/in/saulm24](https://www.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
Master of Science in Computer Science, Georgia Institute of Technology

Graduated 2024
Expected 2027

Professional Experience

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

- Led end-to-end AI and data engineering projects for production services department, including production AI agent deployment, automated data processing systems, and managing stakeholder relationships.
- Received Q1 2025 Spotlight Award within Production Services Department for architecting and deploying AI agent solution that automated ServiceNow data insights, eliminating manual analysis workflows for enterprise stakeholders.
- Built automated data retention algorithms using Oracle PL/SQL and Pentaho, processing 20M+ daily records across multiple databases while ensuring data compliance and referential integrity at enterprise scale.
- Architected 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.
- Managed cross-functional stakeholder relationships for ServiceNow data requirements, designing custom ETL solutions and data ingestion points tailored to business needs.
- 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 5+ 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

CloudScale Finance ETL Pipeline | Python, AWS, BigQuery, Airflow *Spring 2025*

- Built automated financial data pipeline using Python to process daily stock market data from Alpha Vantage API, implementing rate limiting and error handling for reliable data acquisition.
- Deployed cloud ETL architecture with AWS S3, BigQuery, and Airflow orchestration, calculating technical indicators (SMA, RSI, MACD) and automated data quality validation.
- Implemented Infrastructure as Code using Terraform and Docker with GitHub Actions CI/CD, establishing monitoring dashboards for production deployment.

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.