Amrutha Vinayakam

New York, NY | (201) 241-5053 | amrutha513851@gmail.com | GitHub | LinkedIn | Portfolio

SUMMARY

Data Scientist with 4+ years of experience across healthcare analytics, predictive modeling, and AI/ML engineering. Built LLM solutions to operationalize policy-driven eligibility and workflow automation, and financial forecasting models that strengthened revenue planning for healthcare chronic care programs. Spearheaded 5+ ETL pipelines, cutting data preparation time by 40%.

PROFESSIONAL EXPERIENCE

Data Scientist | Food Bank of South Jersey | Pennsauken, NJ **Dec 2025 – Present**

- Architected a centralized Azure data warehouse, designing scalable schemas, ingestion processes, and standardized reporting layers, reducing data discrepancies by 15% and saving 8 hours/week.
- Built an ensemble of Random Forest and Logistic Regression models to segment agencies on demographic and behavioral data, driving a 15% increase in distribution across 200+ partner agencies.
- Deployed an XGBoost demand forecasting model on historical distribution and program data, improving planning accuracy and directing resource allocation decisions for weekly operational reviews.
- Delivered executive-ready Power BI dashboards and governed KPI reporting, enabling weekly and monthly operational reviews by translating program needs into reusable analytics views.
- Aligned Grants, Marketing, and Finance stakeholders on shared KPIs, converting data insights into program strategies adopted in leadership decision-making across the organization.

Artificial Intelligence Engineer | Dreamline AI | Florida, Remote **Jul 2025 – Dec 2025**

- Elevated user behavior prediction by 25% by developing an ensemble of XGBoost and Logistic Regression models on user, property, and policy data, with LLM-based feature enrichment for program rules, while optimizing for latency and cost.
- Increased match accuracy by 10% by segmenting users with K-Means clustering for program alignment, validated with t-tests to ensure statistical significance of key insights.
- Engineered an automated evaluation suite using F1, confusion matrices, and cross-validation, ensuring model reliability across releases; maintained a 95% testing coverage rate.

Graduate Teaching/Technical Assistant | Fordham University GSAS | New York, NY **Sept 2023 – Dec 2024**

- Developed and maintained 5+ ETL data pipelines for academic datasets, reducing data quality issues by 40% by standardizing ingestion, cleaning, and data-validation steps for 30+ students.
- Created 4+ interactive Tableau dashboards for feedback and enrollment KPIs, improving faculty decision-making efficiency by 50% by consolidating metrics into self-serve visual reporting.
- Implemented a Transformers-based classification model, achieving 92% accuracy by fine-tuning Hugging Face models to automate research content tagging and save 10+ hours/month.

Artificial Intelligence Engineer | Radical AI | New York, NY **May 2024 – Oct 2024**

- Accelerated model quality by 25% and real-time processing speed by 30% by iterating on feature pipelines and training workflows using Python, SQL, Scikit-learn, TensorFlow, PyTorch, and Keras.
- Boosted production chatbot response accuracy by 15% by refining NLP pipelines with Hugging Face Transformers and the OpenAI API, including prompt and retrieval tuning.
- Reduced inference latency by 20% by preprocessing and optimizing text pipelines using SpaCy and NLTK for large-scale datasets.
- Optimized pipeline efficiency by 35% by integrating LangChain and LlamaIndex components to increase throughput and support scalable agent workflows.
- Drove up user support engagement by 15% by contributing to AI virtual agent features that expanded real-time assistance coverage.

Data Analyst | KPMG | India **Aug 2020 – Jul 2023**

- Sharpened monthly revenue projections by 20% by building financial forecasting models from historical claims and billing data for chronic care programs.
- Productionized forecasting models (ARIMAX, SARIMAX) and XGBoost in Azure ML Studio for repeatable training and scoring, automating monthly runs to support revenue planning.
- Presented ROI models to cross-functional finance stakeholders, influencing $2M+ investment decisions across 3 digital health pilots.

TECHNICAL SKILLS

Programming: Python (scikit-learn, PyTorch, PySpark, NumPy, pandas, matplotlib, CausalImpact), SQL
Analytics: Machine Learning, A/B Testing, Causal Inference, Predictive Modeling, Regression, Classification, Clustering
Tools: Snowflake, MongoDB, Google BigQuery, Azure, GCP, AWS, Databricks, Power BI, Tableau, Looker, MLOps

EDUCATION

Fordham University, New York — Master of Science in Data Science
Osmania University — Bachelor of Engineering, Electronics and Communication Engineering

CERTIFICATIONS

Google Cloud Associate Data Practitioner (Google, 2026) | Machine Learning with Python – Level 1 (IBM) | Certified Data Scientist (ExcelR)