

Colton Behannon

colton.behannon@gmail.com | linkedin.com/in/coltonbehannon

SUMMARY

Data Scientist with 5+ years of experience building production AI/ML systems and turning complex data into actionable insights. Strong track record of deploying LLM applications, agentic AI systems, and ML models that drive measurable business impact. Expertise in full-stack data science: from infrastructure and MLOps to experimentation, statistical analysis, and translating technical findings into strategic recommendations for technical and executive audiences.

PROFESSIONAL EXPERIENCE

Senior Data Scientist, Generative AI

May 2023 - Present

Equinor, Houston, TX

- Architected and deployed production agentic AI systems using Semantic Kernel with integrated tool calling, providing real-time technology intelligence and reducing analyst research time by 30%. Built end-to-end from API design to monitoring infrastructure.
- Designed enterprise-wide Agentic AI Blueprint defining technical standards for agent orchestration, MCP server development, evaluation frameworks, and guardrails. Framework currently adopted by 4 production projects with company-wide rollout in progress following executive approval.
- Built highly secure multilingual transcription service using OpenAI Whisper optimized for GPU inference with speaker diarization, processing highly confidential data and automating workflows to save 100+ man-hours monthly. Implemented CI/CD pipelines using GitHub Actions, Docker, and Azure Container Registry.
- Developed automated news intelligence pipeline processing 70+ articles monthly using LLM-based metadata extraction and custom ranking algorithms, enabling efficient curation for technology intelligence. Implemented RAG system with vector search (Azure AI Search) and Neo4j knowledge graphs.
- Established monitoring and observability practices for production AI systems, implementing evaluation pipelines and creating dashboards to track model performance, latency, and error rates across distributed infrastructure.
- Collaborated across business units with 10+ stakeholders including engineering, product, and finance teams to deliver data-driven insights and technical strategy. Presented technical findings to VP-level executives and shared AI best practices with 50+ developers at company conference.

Technical Environment: Python, Azure AI Foundry, Semantic Kernel, LangChain/LangGraph, OpenAI Whisper, Arize Phoenix, Kubernetes, Docker, Azure ML, Neo4j, Vector Search, HuggingFace TGI, CI/CD (GitHub Actions)

Data Scientist

December 2021 - April 2023

Perry Homes, Houston, TX

- Built XGBoost-based forecasting models for construction timelines across 1,500+ active projects, improving prediction accuracy by 2.5x and enabling better resource allocation. Leveraged feature engineering on construction stage durations to create interpretable models for operational stakeholders.
- Conducted comprehensive statistical analysis of construction options packages across 90+ communities, identifying pricing optimization opportunities worth \$1M-\$4.5M per quarter through pattern analysis of negotiated pricing data and consumer behavior.
- Designed and implemented automated data pipelines using SQL for field data collection, establishing reliable data infrastructure powering company-wide analytics. Created PowerBI dashboards and SQL views for real-time tracking of construction progress and KPIs.
- Delivered data insights to executive and operational audiences including city presidents and construction managers across 4 regional offices, translating complex analytical findings into actionable business recommendations.

Technical Environment: Python (pandas, scikit-learn, XGBoost), SQL (MSSQL), Azure ML, PowerBI

Statistician

May 2021 - December 2021

Alliantgroup, L.P., Houston, TX

- Conducted 500+ statistical samples for R&D tax credit analysis using sampling theory and hypothesis testing, generating \$3M+ in client benefits while reducing employee interview time by hundreds of hours.
- Refactored legacy R codebase to comply with updated tax code requirements, ensuring continued business operations and maintaining statistical rigor across all analysis workflows.

EDUCATION

M.S. Statistics and Data Science

2021

University of Houston, Houston, TX

BBA, Management Information Systems

2020

University of Houston, Houston, TX

Summa Cum Laude, Dean's List (2017-2020), Academic Excellence Scholarship

TECHNICAL SKILLS

AI/ML & Frameworks: LLM Applications, Agentic AI, Azure AI Foundry, Semantic Kernel, LangChain/LangGraph, HuggingFace Transformers, RAG, Prompt Engineering, Model Evaluation (Arize Phoenix)

Machine Learning: Gradient Boosting (XGBoost), scikit-learn, Statistical Modeling, Time Series Forecasting, Feature Engineering

MLOps & Infrastructure: Kubernetes, Docker, CI/CD (GitHub Actions), Azure ML, LLMOps, Model Monitoring, Distributed Systems

Data Engineering: SQL (MSSQL, PostgreSQL), Data Pipelines, Vector Search (Azure AI Search), Neo4j, ETL/ELT

Programming & Tools: Python (pandas, NumPy, scikit-learn), R, SQL, JavaScript, Git, PowerBI

Analysis & Experimentation: A/B Testing, Statistical Inference, Hypothesis Testing, Causal Inference, Experimental Design, Data Visualization