

540.219.6441
Washington D.C.
ngohlmartin@gmail.com
Portfolio: Martin Ngoh

Martin Ngoh

Senior Data Scientist | AI Engineer

github.com/mngoh
linkedin.com/mngoh

TECHNICAL SKILLS

Languages **Proficient:** Python, PySpark, SQL, GIT, HTML
 Experience: R, SAS, JavaScript, CSS

Tech Stack AWS, Azure, Tableau, Tensorflow, PyTorch, MLFlow, LangChain, numpy, pandas, matplotlib, seaborn, sklearn, Hugging Face, Prophet, OpenAI

EXPERIENCE

Senior Data Scientist **July 2024 — Present**
Accenture Federal Services *Arlington, VA*

- Developed and deployed a forecasting model using Prophet (MAPE: 12%), with self-healing, allowing for accurate prediction of computational load, resulting in more than \$1 million in annual savings.
- Built a novel, fully automated, modeling pipeline in AWS using Lambda, providing real-time model results through QuickSight and Athena, and daily self-healing model capabilities using batch transform jobs.
- Identified and recommended a novel approach to using AI for fraud detection. Independently designed and delivered a curated fraud detection model using XGBoost (recall: 85%). The results were delivered to the client and future work was assigned due to this initiative.
- Mentor junior data scientists, providing code reviews and coaching sessions to help build their confidence in delivering quality work while uplifting the team's performance as a whole.

Data Scientist **October 2021 — July 2024**
Deloitte Consulting *Arlington, VA*

- Strategized a road map with future AI initiatives for the client, helping secure a \$30 million contract over 3 years.
- Boosted fraud detection by 8% by implementing Bayesian optimization (BayesSearchCV) to tune an ensemble model of neural network, elastic net and gradient boosting models.
- Refactored ML models to integrate with MLflow, enabling centralized model tracking and cutting documentation by 90%. This approach allowed for rapid experimentation, leading to \$100 million more in savings for the client.
- Automated a 12-hour program using Linux, reducing turnaround time from 3 days to 2 hours by leveraging cron jobs. Led a team training program for 20 colleagues, which resulted in a 90% increase in efficiency.

PROJECTS

GenAI RAG Model - JournalBot | Python, OpenAI, Whisper, Streamlit, NLTK **2026**

- Engineered an end-to-end RAG Pipeline to transform 10+ hours of raw voice recordings into a searchable personal knowledge base using Python and OpenAI's Whisper for transcription.
- Optimized data retrieval by chunking transcripts into a ChromaDB vector store; implemented vector embeddings and cosine similarity to enable efficient, context-aware query results.
- Developed an interactive UI using Streamlit to allow for prompts (e.g., "What are my thoughts on cookies?") with low-latency response times.
- Performed sentiment analysis, leveraging NLTK to provide metadata and visual plots on emotional tone and recurring themes within journal entries.

EDUCATION

M.S. Business Analytics, Georgetown University (Varsity Men's Soccer Team) 2022
B.S. Supply Chain Management and Analytics, VCU (Varsity Men's Soccer Team) 2020

EXTRACURRICULAR

- Selected as a keynote speaker at Georgetown University's 2024 Black Alumni Summit; spoke on a panel of five AI experts discussing the implications of AI for the Black Community, to an audience of over 200.