Matthew Martin

Email: matthewmartin117@gmail.com

Website: matthewmartin117.github.io — GitHub: github.com/matthewmartin117

Summary: AI Engineer with a background in **Computer Science and Biology** and hands-on experience designing and deploying intelligent systems. Skilled in building **AI-powered agents** and machine learning models for **workflow automation** and data analysis in complex, regulated domains. Passionate about leveraging **LLMs** and multimodal data to solve challenges in compliance, resource optimization, and operational efficiency.

Education

University of Colorado Boulder — B.S. **Computer Science**, GPA: 4.0/4.0, Expected Dec 2025

– Dean's List, CSPB Orientation Leader

James Madison University — B.S. **Biology**, Graduated Jul 2021

– Dean's List

Technical Skills

AI/ML & Data Science: Python (**NumPy, Pandas, Scikit-learn**), PyTorch, TensorFlow/Keras, HuggingFace (**LLMs**), NLP, **Anomaly Detection**

Programming Cloud: Java, JavaScript/React, Spring Boot, **AWS** (S3, EC2, SageMaker), Docker, Git/GitHub, **APIs**, Microservices

Data Systems: SQL (PostgreSQL), **ETL ML pipelines**, Multimodal Data Fusion

Work Experience

Data Engineer Intern — **AstraZeneca**, May 2025 – Aug 2025

– Built and maintained **ETL pipelines** for experimental datasets using SQL Server.
– Developed a C#/.NET application to automate data ingestion, improving pipeline reliability.
– Partnered with **cross-functional data science teams** in **Agile sprints** to translate business needs into technical requirements for predictive modeling workflows.

Sampling Coordinator — **United States Pharmacopeia**, Mar 2022 – Jan 2023

– Managed pharmaceutical sample logging and shipments via **Oracle LIMS**, ensuring regulatory compliance.

Clinical Research Scientist — **Curative Health**, Aug 2021 – Dec 2021

– Processed large-scale **COVID-19 PCR datasets** with high data integrity.
– Developed methods to ensure accuracy and compliance in **clinical data pipelines**.

Projects

Intelligent Agent for Document Compliance — Java, Spring Boot, LLMs, Docker, AWS

– Architected and deployed a secure **API** to power an intelligent document processing system.
– Built and integrated an **AI agent** using a HuggingFace **LLM** to automatically detect and handle non-compliant sensitive information, supporting audit and compliance workflows.
– Implemented **controller-level CRUD operations** for managing and tracking document status.

Heart Disease Prediction — Python, Scikit-learn, Pandas

– Conducted **EDA** and feature engineering on 400K CDC survey responses to identify key risk factors.
– Trained and evaluated multiple models, including Random Forest and Gradient Boosting, to predict patient risk.
– Achieved approximately **90% recall** for positive cases, creating a valuable tool for identifying at-risk individuals and optimizing resource allocation.

Pneumonia Detection from Chest X-Rays — Python, TensorFlow, Keras

– Developed a deep learning model to analyze **multimodal data** (X-ray images) for disease detection.
– Implemented a robust data pipeline involving preprocessing, class weighting, and augmentation to handle imbalanced data.
– Developed CNN and transfer learning models achieving 0.92 recall, demonstrating strong model validation skills.