Matthew Martin

Phone: (301)-820-5744 | **Email:** matthewmartin117@gmail.com

Education

The University of Colorado Boulder — Bachelor of Science in Computer Science
GPA: **4.00 / 4.00** — Expected Graduation: Dec 2025
Awards: Dean's List Student, CSPB Orientation Leader

James Madison University — Bachelor of Science in Biology
Graduated: Jul 2021

Technical Skills

AI & Machine Learning: TensorFlow, PyTorch, Keras, HuggingFace, Scikit-learn, Pandas, NumPy
Programming Languages: Python, SQL
Tools & Platforms: Git, AWS (S3, Lambda, EC2), Jupyter Notebook,Copilot, GitHub
Concepts: Deep Learning (CNNs, Transfer Learning), NLP, Computer Vision, Supervised & Unsupervised Learning, ETL Pipelines, Data Modeling, MLOps, Responsible AI

Work Experience

- ***Software Engineering Intern*** ***AstraZeneca*** *May 2025 - August 2025*
  - Designed and deployed a scalable data ingestion & preprocessing pipeline migrating experimental datasets from Excel into SQL Server, supporting ML-driven biomedical research.
  - Implemented **schema normalization and validation** for large datasets, ensuring high-quality inputs for predictive modeling.
  - Partnered with
  - data scientists to align pipelines with clinical research use cases.
  - Delivered reproducible solutions through code reviews, demos, and documentation.
- ***Sampling Coordinator*** ***United States Pharmacopeia*** *March 2022 - January 2023*
  - Managed **pharmaceutical data collection and logging** using Oracle systems, ensuring accuracy and integrity across large datasets.
  - Coordinated **global logistics of regulated materials**, requiring precision and compliance with data-driven processes.
- ***Research Associate*** ***Curative Health*** *August - December 2021*
  - **Processed and analyzed** large-scale COVID-19 PCR datasets, improving variant detection and supporting public health decision-making during the pandemic.
  - Developed methods to ensure **data quality at scale**, enabling accurate variant monitoring.

Projects

- **Pneumonia Detection from Chest X-Rays** — *Deep Learning (TensorFlow/Keras)*
  - Built a CNN-based diagnostic assistant (TensorFlow/Keras, VGG16 transfer learning).
  - Applied class weighting & image augmentation to mitigate imbalance, achieving **86% accuracy / 0.92 recall**.
- **Heart Disease Prediction** — *Supervised Learning (Scikit-learn)*
  - Engineered an end-to-end ML pipeline for patient risk stratification using **400k+ records**.
  - Applied feature engineering and ensemble models (**Random Forest, Gradient Boost**), achieving **90% recall** for at-risk patients.
  - Showcased scalable applications of AI in preventative care.
- **Breast Cancer Gene Expression Clustering -** *PCA, K-Means, Unsupervised Learning*
  - Processed a high-dimensional dataset of **54,000+ gene features** by applying **PCA** for dimensionality reduction and **K-Means clustering** to identify distinct biological patterns and patient subtypes across samples.