

# VU TRONG CHAU

Manor, TX | +1 (904) 820-9298 | [chautrongvu@gmail.com](mailto:chautrongvu@gmail.com) | [Vu Chau | LinkedIn](#) | [Vu Trong Chau - Portfolio](#)

## SUMMARY

Data Scientist & Machine Learning Engineer with hands-on experience building LLM applications, RAG pipelines, predictive models, and end-to-end MLOps systems. Skilled in developing and deploying production-grade AI solutions across AWS, GCP, and Azure. Strong background in data engineering, cloud-native architecture, vector search, analytics, and real-time model inference. Demonstrated ability to transform complex datasets into actionable insights and scalable ML systems.

## TECHNICAL SKILLS

**Programming Languages:** Python, R, C/C++, JavaScript, SQL

**Machine Learning & AI:** LLMs, RAG, NLP, Deep Learning, CNNs, RNNs, XGBoost, SVM

**Frameworks:** PyTorch, TensorFlow, Scikit-learn, LangChain, HuggingFace

**MLOps & Cloud:** Docker, Kubernetes, CI/CD, MLflow, Airflow, FastAPI, AWS (SageMaker), GCP (Vertex AI), Azure ML

**Data Engineering:** Pandas, NumPy, SQL, D3.js, Tableau, ETL pipelines, Data Warehousing

## PROJECTS

### Healthcare Chatbot

Sep 2025 – present

- Designed and deployed an LLM-powered healthcare Q&A chatbot using RAG on 250,000+ medical records.
- Built an automated MLOps pipeline using CI/CD, Docker, and FastAPI, reducing deployment cycle time by 40%.
- Integrated FAISS vector search and optimized prompts, improving retrieval accuracy by 25%.

### Threat Detection using Machine Learning

Jan 2025 – July 2025

- Developed multiple ML models (Logistic Regression, XGBoost, Random Forest, Naive Bayes), achieving 92% accuracy on 72,000+ labeled texts.
- Engineered semantic, syntactic, and sentiment features, increasing classifier performance by 15%.
- Built a real-time moderation interface using Flask, generating severity scores and reducing review time by 40%.

### Global Population Prediction

Jan 2025 – May 2025

- Forecasted population trends (1960–2023) for 200+ countries using ML models achieving <2% MAE.
- Built interactive dashboards (choropleth, line charts, bar charts) to improve researcher usability and insight discovery.
- Implemented scalable data-refresh pipelines using Python and D3.js, reducing manual update effort by 60%.

### Sleep Quality Prediction

Sep 2024 – Dec 2024

- Developed machine learning models (Logistic Regression, Random Forest, CNNs, RNNs) to predict sleep quality from 70,000+ lifestyle and biometric records.
- Built ML models predicting sleep quality using lifestyle and biometric data, improving precision by 20% over baseline
- Created a real-time prediction interface via Flask, providing personalized recommendations to users.

## EDUCATION

### Troy University

Troy, AL

**Master of Science in Computer Science - Artificial Intelligence** | GPA: 3.5

July 2025

Key Courses: Analysis of Algorithms, Computer Architecture, Machine Learning, Advanced Artificial Intelligence, Data Visualization.

### University of Sunderland

Sunderland, UK

**Bachelor of Engineering in Electronic and Electrical Engineering** | UK 2:1 Honours (3.5 GPA equivalent)

July 2021

Key Courses: Embedded systems, Electronic Circuits and Devices, Electrical Power, Electronic Systems, Manufacturing System Design.

## CERTIFICATIONS

LLM Application Engineering and Development | **Simplilearn** | [Link](#)

Oct 2025

Data Science Methodology | **IBM** | [Link](#)

Sep 2025

Generative AI with Large Language Models | **DeepLearning.AI** | [Link](#)

Sep 2025